

FUNDAMENTOS DE

Tercera edición

SQL

Cubre:
SQL:2006 ANSI/ISO standard • SQL/XML •
La última versión de los programas RDBMS.

Andy Oppel
Robert Sheldon



www.detodoprogramacion.com

Libros Universitarios, juegos, Revistas, Cursos, Software, Sistemas Operativos, Antivirus y más ... Gratis para el Conocimiento...!

www.detodoprogramas.com

Visítanos y compruébalo

DETODOPROGRAMAS.COM



Material para los amantes de la Programación Java, C/C++/C#, Visual.Net, SQL, Python, Javascript, Oracle, Algoritmos, CSS, Desarrollo Web, Joomla, jquery, Ajax y Mucho Mas...

www.detodoprogramacion.com

Visitanos



Fundamentos de **SQL**

Tercera edición

Andy Oppel
Robert Sheldon

Traducción

Carlos Fabián Jiménez Castillo

Traductor profesional



MÉXICO • BOGOTÁ • BUENOS AIRES • CARACAS • GUATEMALA • MADRID • NUEVA YORK
SAN JUAN • SANTIAGO • SÃO PAULO • AUCKLAND • LONDRES • MILÁN • MONTREAL
NUEVA DELHI • SAN FRANCISCO • SINGAPUR • ST. LOUIS • SIDNEY • TORONTO

www.detodoprogramacion.com

Director editorial: Fernando Castellanos Rodríguez
Editor: Miguel Ángel Luna Ponce
Supervisor de producción: Zeferino García García

FUNDAMENTOS DE SQL
Tercera edición

Prohibida la reproducción total o parcial de esta obra,
por cualquier medio, sin la autorización escrita del editor.



Educación

DERECHOS RESERVADOS © 2010, respecto a la primera edición en español por
McGRAW-HILL INTERAMERICANA EDITORES, S.A. DE C.V.

A Subsidiary of The McGraw-Hill Companies, Inc.

Corporativo Punta Santa Fe,
Prolongación Paseo de la Reforma 1015, Torre A
Piso 17, Colonia Desarrollo Santa Fe,
Delegación Álvaro Obregón
C.P. 01376, México, D.F.

Miembro de la Cámara Nacional de la Industria Editorial Mexicana, Reg. Núm. 736

ISBN: 978-607-15-0251-3

Translated from the 3rd English edition of

SQL: A Beginner's Guide

By: Andrew Oppel and Robert Sheldon

Copyright © 2009 by The McGraw-Hill Companies. All rights reserved.

ISBN: 978-0-07-154864-9

1234567890

109876543210

Impreso en México

Printed in Mexico

Acerca de los autores

Andrew (Andy) J. Oppel es graduado de la Latin School de Maryland y la Transylvania University (Lexington, Kentucky) donde obtuvo un grado en ciencias de la computación en 1974. Desde entonces ha sido empleado continuamente en una amplia variedad de puestos sobre tecnologías de la información, incluyendo programador, programador/analista, arquitecto en sistemas, administrador de proyectos, administrador senior de base de datos, administrador del grupo de bases de datos, asesor, diseñador de bases de datos, modelador de datos y arquitecto de datos. Además, ha sido instructor de medio tiempo para un campus de la Universidad de California (Berkeley) durante más de 20 años, y recibió el premio de Instructor Honorario durante el año 2000. Su labor en la enseñanza incluye el desarrollo de tres cursos para el campus de la Universidad de California, “Conceptos de Sistemas de Administración de Bases de Datos”, “Introducción a los Sistemas de Administración de Bases de Datos Relacionales”, y “Modelado de Datos y Diseño de Bases de Datos”. También obtuvo la certificación Oracle 9i Database Associate en 2003. Actualmente se encuentra empleado como modelador de datos senior para Blue Shield de California. Muy aparte de los sistemas de cómputo, Andy disfruta la música (la guitarra y el canto), el radio amateur (vicedirector de la Pacific Division, American Radio Relay League) y el fútbol (instructor de árbitros, U.S. Soccer).

Andy ha diseñado e implementado cientos de bases de datos para un amplio rango de aplicaciones, incluyendo investigación médica, para la banca, seguros, fabricación de instrumentaria, telecomunicaciones, comunicaciones inalámbricas y recursos humanos. Es el autor de *Databases Demystified* (McGraw-Hill/Osborne, 2004) y *SQL Demystified* (McGraw-Hill/Osborne, 2005). Su experiencia en productos de bases de datos incluye IMS, DB2, Sybase, Microsoft SQL Server, Microsoft Access, MySQL y Oracle (versiones 7, 8, 8i, 9i y 10g).

Robert Sheldon ha trabajado como asesor y escritor técnico durante muchos años. Como asesor, ha administrado el desarrollo y mantenimiento de aplicaciones basadas en Web y aplicaciones servidor-cliente y las bases de datos que soportan a esas aplicaciones. Ha diseñado e implementado diferentes bases de datos de Access y SQL y ha utilizado SQL para construir bases de datos, crear y modificar objetos de las bases de datos, consultar y modificar los datos, y para solucionar problemas relacionados con los datos y con los sistemas. Robert también ha escrito o coescrito ocho libros sobre diferentes tecnologías de red y de servidor, uno de los cuales recibió un certificado al mérito del Puget Sound Chapter de la Society for Technical Communication. Además, dos de los libros que Robert ha escrito se enfocan ex-

clusivamente al diseño e implementación de SQL Server. Robert también ha escrito y editado una gran cantidad de material relacionado con las bases de datos SQL y otras tecnologías de computación. Su trabajo en escritura no sólo incluye material sobre la industria de la computación (también ha escrito un poco de todo, desde artículos de noticias hasta trabajos en documentación legal) y ha recibido dos premios de la Colorado Press Association.

Acerca del revisor técnico

James Seymour es graduado de la universidad de North Carolina en Chapel Hill con grado en historia y ciencias políticas, y de la universidad de Kentucky con una maestría en historia. Se involucró por primera vez con tecnología computacional en 1965 con el entorno de la computadora central en North Carolina. En el ejército de Estados Unidos durante la guerra de Vietnam perteneció al pequeño equipo que trabajó con la puesta en marcha de la computadora central del Pentágono para diferentes escenarios de estrategia militar.

Desde 1972 está involucrado en diferentes entornos de computación con el segundo proyecto punto de venta y control de inventarios en la industria de venta al público, programas analíticos e iniciativas de bases de datos para industrias de seguros y ayuda social, la puesta en marcha de controles de pérdida, y otros proyectos de control de inventario y rastreo de ventas a través de diferentes industrias.

De 1987 a 1995, James fue instructor de manejo de bases de datos en el sistema de universidades comunitarias en el estado de Kentucky. Con esta capacidad creó los primeros cursos de manejo de bases de datos y de programación en C en el estado de Kentucky y ayudó a las entidades tanto públicas como privadas con sus necesidades urgentes de entrenamiento, incluyendo la programación de sistemas guía para los misiles crucero para Tormenta del Desierto.

Antes de 1985 fue administrador de sistemas, administrador de redes, programador y administrador de bases de datos. A partir de 1985, James ha sido administrador senior de bases de datos y ha trabajado principalmente con DBMS de DB2 y Oracle en múltiples plataformas, incluyendo SQL Server desde la versión 7.0. Actualmente es el administrador senior de bases de datos y arquitecto de datos para una compañía del Fortune 100 supervisando proyectos mayores en Estados Unidos, Canadá y el Reino Unido.

Contenido

ACERCA DE LOS AUTORES	iii
AGRADECIMIENTOS	xi
INTRODUCCIÓN	xi

Parte 1 Bases de datos relacionales y SQL

1 Introducción a las bases de datos relacionales y a SQL	3
Entienda las bases de datos relacionales	4
El modelo relacional	5
Aprenda acerca de SQL	15
La evolución de SQL	15
Tipos de instrucciones de SQL	18
Tipos de ejecución	19
El estándar SQL frente a las implementaciones de producto	21
2 Trabajo con el entorno SQL	29
Entienda el entorno SQL	30
Entienda los catálogos SQL	32
Esquemas	34
Objetos de esquema	35
¿Qué es una base de datos?	37
Nombrado de objetos en un entorno SQL	40
Nombres calificados	41
Creación de un esquema	42
Creación de una base de datos	44

3 Creación y modificación de tablas.....	49
Creación de tablas en SQL.....	50
Especificación de los tipos de datos en una columna.....	54
Tipos de datos de cadena.....	55
Tipos de datos numéricos.....	57
Tipos de datos de fecha y hora.....	58
Tipo de datos de intervalo.....	60
Tipo de datos booleanos.....	61
Utilice tipos de datos SQL.....	62
Creación de tipos definidos por el usuario.....	63
Especificación de los valores predeterminados en una columna.....	64
Eliminación de tablas en SQL.....	69
4 Implementación de la integridad de datos.....	73
Entienda las restricciones de integridad.....	74
Utilice restricciones NOT NULL.....	76
Añada restricciones UNIQUE.....	77
Añada restricciones PRIMARY KEY.....	79
Añada restricciones FOREIGN KEY.....	83
La cláusula MATCH.....	88
La cláusula <acción referencial desencadenada>.....	89
Defina restricciones CHECK.....	95
Defina afirmaciones.....	97
Creación de dominios y restricciones de dominio.....	98
5 Creación de vistas en SQL.....	103
Añada vistas a la base de datos.....	104
Definición de vistas de SQL.....	108
Creación de vistas actualizables.....	114
Utilice la cláusula WITH CHECK OPTION.....	116
Eliminación de vistas de la base de datos.....	117
6 Gestión de seguridad en la base de datos.....	123
Entienda el modelo de seguridad de SQL.....	124
Sesiones SQL.....	126
Acceda a objetos de base de datos.....	128
Creación y eliminación de roles.....	130
Otorgue y revoque privilegios.....	131
Revoque privilegios.....	135
Otorgue y revoque roles.....	137
Revoque roles.....	138

Parte II Acceso y modificación de datos

7 Consulta de datos de SQL	145
Utilice la instrucción SELECT para la recuperación de datos.	146
La cláusula SELECT y la cláusula FROM	147
Utilice la cláusula WHERE para definir condiciones de búsqueda	152
Defina la cláusula WHERE	156
Utilice la cláusula GROUP BY para agrupar los resultados de una consulta	159
Utilice la cláusula HAVING para especificar un grupo de condiciones de búsqueda	164
Utilice la cláusula ORDER BY para ordenar los resultados de una consulta	166
8 Modificar datos SQL	175
Insertar datos SQL	176
Insertar valores desde una instrucción SELECT	180
Actualizar datos SQL	182
Actualizar valores desde una instrucción SELECT	185
Eliminar datos SQL	186
9 Utilizar predicados	193
Comparar datos SQL	194
Utilizar el predicado BETWEEN	199
Arrojar valores nulos	200
Arrojar valores similares	203
Hacer referencia a fuentes adicionales de datos	209
Utilizar el predicado IN	209
Utilizar el predicado EXISTS	213
Determinar la cantidad de predicados de comparación	216
Utilizar los predicados SOME y ANY	216
Utilizar el predicado ALL	218
10 Trabajar con funciones y expresiones de valor	225
Utilizar funciones Set	226
Utilizar la función COUNT	227
Utilizar las funciones MAX y MIN	229
Utilizar la función SUM	231
Utilizar la función AVG	232
Utilizar funciones de valor	232
Trabajar con funciones de valor de cadena	233
Trabajar con funciones de valor de fecha y hora	236
Utilizar expresiones de valor	238
Trabajar con expresiones de valor numéricas	238

Utilizar la expresión de valor CASE.....	241
Utilizar la expresión de valor CAST.....	244
Utilizar valores especiales	245
11 Acceder a múltiples tablas.....	253
Realizar operaciones básicas join	254
Utilizar nombres de correlación	257
Crear operaciones join con más de dos tablas	258
Crear la operación cross join	259
Crear la operación self-join.....	260
Unir tablas con nombres de columna compartidos	261
Crear el método join natural	262
Crear el método join de columna nombrada.....	263
Utilizar el método join de condición	263
Crear la inner join	264
Crear la outer join	266
Realizar operaciones de unión	269
12 Utilizar subconsultas para acceder y modificar datos.....	277
Crear subconsultas que arrojen múltiples filas	278
Utilizar el predicado IN	279
Utilizar el predicado EXISTS.....	281
Utilizar predicados de comparación cuantificados.....	282
Crear subconsultas que arrojen un solo valor	283
Trabajar con subconsultas correlacionadas	284
Utilizar subconsultas anidadas.....	286
Utilizar subconsultas para modificar datos	288
Utilizar subconsultas para insertar datos	288
Utilizar subconsultas para actualizar datos.....	290
Utilizar subconsultas para eliminar datos.....	291
 Parte III Acceso avanzado a los datos	
13 Crear rutinas invocadas por SQL.....	299
Entender las rutinas invocadas por SQL	300
Procedimientos y funciones invocadas por SQL	301
Trabajar con la sintaxis básica	302
Crear procedimientos invocados por SQL.....	303
Invocar procedimientos invocados por SQL.....	305
Agregar parámetros de entrada a sus procedimientos	306
Utilizar procedimientos para modificar datos.....	309
Agregar variables locales a sus procedimientos.....	311

Trabajar con instrucciones de control	313
Crear instrucciones compuestas	313
Crear instrucciones condicionales	314
Crear instrucciones de repetición	316
Agregar parámetros de salida a sus procedimientos	320
Crear funciones invocadas por SQL	321
14 Crear activadores SQL.....	329
Entender los activadores SQL	330
Contexto de ejecución del activador	331
Crear activadores SQL	333
Referenciar valores antiguos y nuevos	334
Quitar activadores SQL	335
Crear activadores de inserción	336
Crear activadores de actualización	338
Crear activadores de eliminación	343
15 Utilizar cursores SQL.....	351
Entender los cursores SQL	352
Declarar y abrir cursores SQL	353
Declarar un cursor	355
Trabajar con elementos opcionales de la sintaxis	356
Crear una instrucción de cursor	360
Abrir y cerrar un cursor	363
Recuperar datos desde un cursor	363
Utilizar instrucciones UPDATE y DELETE posicionadas	368
Utilizar la instrucción UPDATE posicionada	369
Utilizar la instrucción DELETE posicionada	370
16 Manejar transacciones SQL	377
Entender las transacciones SQL	378
Configurar las propiedades de la transacción	381
Especificar un nivel de aislamiento	382
Especificar un tamaño de diagnóstico	387
Crear una instrucción SET TRANSACTION	388
Iniciar una transacción	389
Determinar el aplazamiento de una restricción	390
Crear puntos de recuperación en una transacción	392
Liberar un punto de recuperación	394
Finalizar una transacción	395
Completar una transacción	395
Reinvertir una transacción	396

17 Acceder a datos SQL desde un programa host	403
Invocar SQL directamente	404
Incrustar instrucciones SQL en el programa	406
Crear una instrucción SQL incrustada	407
Utilizar variables host en las instrucciones SQL	408
Recuperar datos SQL	411
Manejo de errores	413
Crear módulos cliente de SQL	417
Definir módulos cliente de SQL	418
Utilizar una interfaz de nivel de llamada de SQL	419
Asignar indicadores	421
Ejecutar instrucciones SQL	423
Trabajar con variables host	424
Recuperar datos SQL	426
18 Trabajar con datos XML	433
Aprender los conceptos básicos de XML	434
Aprender acerca de SQL/XML	437
El tipo de datos XML	437
Funciones SQL/XML	439
Reglas de trazado de SQL/XML	441
 Parte IV Apéndices	
A Respuestas a los autoexámenes	449
B Palabras clave de SQL:2006	491
Palabras clave reservadas de SQL	492
Palabras clave no reservadas de SQL	494
C Código SQL utilizado en los ejercicios Pruebe esto.	497
Código SQL por cada ejercicio	498
La base de datos INVENTARIO	514
Índice	519

Agradecimientos

Hubo mucha gente involucrada en el desarrollo de *Fundamentos de SQL, tercera edición*, muchos de los cuales ni siquiera conozco por nombre. Primero que nada, a los editores y el personal en McGraw-Hill Professional que proporcionaron incontables horas de soporte para este proyecto. Me gustaría agradecer especialmente a la editora Jane K. Brownlow, a la coordinadora de compras Jennifer Housh y a Wilson Drozdowski quien cubrió brevemente a Jennifer, y a todas las personas con quienes individualmente tuve contacto directo a través de todo el proceso de escritura y edición. Sus comentarios y sugerencias, así como sus rápidas y exactas respuestas a muchas de mis preguntas, hicieron que las tareas de escritura se realizaran sin ningún problema; su trabajo tras bambalinas mantuvo al proyecto completo avanzando suavemente. También me gustaría agradecer al corrector de estilo y todos los demás editores, correctores de pruebas, indiceros, diseñadores, ilustradores y otros participantes cuyos nombres desconozco. Un agradecimiento muy especial va dedicado para mi amigo y ex colega Jim Seymour, revisor técnico, por su atención a los detalles y sus aportaciones de mucha ayuda a través de todo el proceso de edición. También me gustaría agradecer el trabajo de Robert Sheldon, autor de las primeras dos ediciones, cuyo excelente trabajo en escritura hizo que las revisiones necesarias para esta edición fueran mucho más fáciles de lograr. Finalmente, quiero agradecer a mi familia por su apoyo y comprensión para poder organizar un horario para la escritura y creación de este libro en mi ya muy ocupada vida.

Andy Oppel

Introducción

Las bases de datos relacionales se han convertido en el mecanismo de almacenamiento de datos más común para las aplicaciones computacionales modernas. Los lenguajes de programación como Java, C y COBOL, y los lenguajes interpretados de programación como Perl, VBScript y JavaScript muy a menudo acceden a las fuentes de datos para poder recuperar o modificar los datos. Muchas de estas fuentes de datos son administradas a través de un sistema de administración de bases de datos relacionales (RDBMS), como Oracle, Microsoft SQL Server, MySQL y DB2, que tiene como base el Lenguaje de Consulta Estructurado (SQL) para crear y alterar los objetos de la base de datos, agregar datos y eliminarlos de la base de datos, modificar datos que han sido agregados a esa base de datos y, por supuesto, recuperar datos almacenados en la base de datos para su desplegado y procesamiento.

SQL es el lenguaje más ampliamente implementado para las bases de datos relacionales. De la misma manera que las matemáticas son el lenguaje de la ciencia, SQL es el lenguaje de las bases de datos relacionales. SQL no solamente permite administrar los datos dentro de la base de datos, sino también manejar la base de datos en sí.

Utilizando las instrucciones SQL, es posible acceder a una base de datos SQL directamente al utilizar una aplicación cliente interactiva o a través de un lenguaje de programación de aplicación o

lenguaje interpretado. Sin importar cuál sea el método que se utilice para acceder a una fuente de datos, es obligatoria una buena base acerca de cómo escribir instrucciones SQL para poder acceder a los datos relacionales. *Fundamentos de SQL, tercera edición*, le proporciona esa base. Se describen todos los tipos de instrucciones que soporta SQL y se explica cómo son utilizadas para administrar bases de datos y sus datos. A través del trabajo con este libro, usted construirá fuertes cimientos en SQL básico y obtendrá un profundo entendimiento de cómo utilizar SQL para acceder a los datos en su base de datos relacional.

Esta tercera edición ha sido actualizada para incluir las disposiciones del estándar ISO SQL:2006 además de todos sus errores corregidos que se publicaron en 2007. El capítulo 18 ha sido incluido para cubrir SQL/XML, que fue agregado al estándar SQL en 2006. Además, las instrucciones SQL han sido reformateadas y todos los nombres de objeto de la base de datos han sido convertidos a mayúsculas para hacer más fácil su lectura y conversión, a través de la amplia variedad de los productos RDBMS disponibles comercialmente.

Quién deberá leer este libro

Fundamentos de SQL está recomendado para cualquiera que busque construir una base en programación de SQL basado en el estándar ISO SQL:2006. El libro está diseñado específicamente para aquellos que son nuevos, o relativamente nuevos en SQL; sin embargo, aquellos que necesitan refrescar sus conocimientos en SQL también encontrarán beneficios en este libro. Ya sea que usted sea un programador experimentado, tenga cierta experiencia en el desarrollo Web, sea un administrador de bases de datos o sea completamente nuevo en programación y bases de datos *Fundamentos de SQL* le proporcionará fuertes bases que serán útiles para cualquiera que desee aprender más acerca de SQL. De hecho, cualquiera de los siguientes ejemplos de personas encontrará útil este libro cuando intente comprender y utilizar SQL:

- El novato que no tenga conocimiento acerca del diseño de bases de datos y programación SQL
- El analista o administrador que quiera una mejor comprensión de cómo implementar y acceder a las bases de datos SQL
- El administrador de bases de datos que quiera aprender más acerca de programación
- El profesional de soporte técnico o ingeniero de pruebas/control de calidad quien realiza consultas *ad hoc* en alguna fuente de datos SQL
- El desarrollador Web que escribe aplicaciones que deberán acceder a bases de datos SQL
- El programador del lenguaje de tercera generación (3GL) que incrusta SQL dentro del código fuente de una aplicación
- Cualquier otro individuo que desee aprender acerca de cómo escribir código SQL que pueda ser utilizado para crear y acceder a bases de datos dentro de algún RDBMS

Cualquiera que sea la categoría en la que se encuentre, un punto importante a recordar es que el libro está enfocado a cualquiera que quiera aprender SQL estándar, y no una versión específica de un producto o de este lenguaje. La ventaja de esto es que es posible tomar las habilidades aprendidas en este libro y aplicarlas a situaciones del mundo real, sin estar limitado a los estándares de un solo producto.

Aún necesitará, por supuesto, conocer cómo implementa SQL el producto en el que está trabajando; pero con la base proporcionada por este libro, será capaz de moverse entre un RDBMS y otro

y aún tener una comprensión básica de cómo se utiliza SQL. Como resultado, este libro es una herramienta útil para cualquier persona que no sepa utilizar las bases de datos basadas en SQL, sin importar el producto que utilice. Los programadores de SQL necesitarán solamente adaptar su conocimiento al RDBMS específico.

Qué contenido cubre el libro

Fundamentos de SQL está dividido en tres partes. La parte I introduce al lector a los conceptos básicos de SQL y explica cómo crear objetos dentro de la base de datos. La parte II proporciona una base acerca de cómo recuperar datos desde una base de datos y modificar (agregar, cambiar y eliminar) esos datos que se encuentran almacenados en la base de datos. La parte III proporciona información acerca de técnicas avanzadas de acceso de datos que permiten expandir lo que se aprendió en las partes I y II. Adicional a estas tres partes, *Fundamentos de SQL* contiene apéndices que incluyen material de referencia para la información presentada en las tres partes.

Descripción de los contenidos del libro

El siguiente resumen describe los contenidos del libro y muestra cómo el libro se encuentra separado en capítulos que se enfocan en determinadas tareas.

PARTE I: Bases de datos relacionales y SQL

Capítulo 1: Introducción a las bases de datos relacionales y a SQL

En este capítulo se encuentra la introducción a las bases de datos relacionales y al modelo relacional, que forma la base para SQL. También se proporcionará un vistazo general a SQL y cómo se relaciona con los RDBMS.

Capítulo 2: Trabajo con el entorno SQL

Este capítulo describe los componentes que conforman el entorno SQL. También se encontrará una introducción a los objetos que conforman un esquema, y se aprenderá cómo crear un esquema dentro del entorno SQL. Se presentará el concepto de crear un objeto de la base de datos en una implementación SQL que soporte la creación de objetos de base de datos.

Capítulo 3: Creación y modificación de tablas

En este capítulo se aprenderá cómo crear tablas SQL, especificar tipos de datos de columna, crear tipos definidos por el usuario y especificar valores de columna por defecto. Se aprenderá cómo alterar una definición de tabla y eliminar esa definición de la base de datos.

Capítulo 4: Implementación de la integridad de datos

Este capítulo explica cómo se utilizan las restricciones de integridad para reforzar la integridad de los datos en las tablas SQL. El capítulo incluye información sobre restricciones relacionadas con las tablas, afirmaciones y restricciones de dominio. Se aprenderá cómo crear restricciones NOT NULL, UNIQUE, PRIMARY KEY, FOREIGN KEY y CHECK.

Capítulo 5: Creación de vistas en SQL

En este capítulo se aprenderá cómo agregar vistas a las bases de datos SQL. También se aprenderá cómo crear vistas actualizables y cómo abandonar vistas de la base de datos.

Capítulo 6: Gestión de seguridad en la base de datos

En este capítulo se introducirá el modelo de seguridad de SQL y se aprenderá cómo se definen los identificadores de autorización dentro del contexto de una sesión. Luego, se aprenderá cómo crear y eliminar roles, otorgar y revocar privilegios, y otorgar y revocar roles.

PARTE II Acceso y modificación de datos

La parte II explica cómo acceder y modificar los datos en una base de datos SQL. También se aprenderá cómo utilizar predicados, funciones y valores de expresión para manipular esos datos. Adicionalmente, la parte II describe cómo unir tablas y utilizar subconsultas para acceder a los datos en múltiples tablas.

Capítulo 7: Consulta de datos de SQL

Este capítulo describe los componentes básicos de la instrucción `SELECT` y cómo la instrucción es utilizada para recuperar datos desde una base de datos SQL. Se aprenderá cómo definir cada cláusula que pueda ser incluida en la instrucción `SELECT` y cómo esas cláusulas son procesadas cuando se consulta una base de datos.

Capítulo 8: Modificar datos SQL

En este capítulo se aprenderá cómo modificar los datos en una base de datos SQL. Específicamente, se aprenderá cómo insertar datos, actualizar datos y eliminar datos. El capítulo repasa cada componente de las instrucciones SQL que permiten realizar esas modificaciones a los datos.

Capítulo 9: Utilizar predicados

En este capítulo se aprenderá cómo utilizar predicados para comparar datos SQL, arrojar valores nulos, arrojar valores similares, hacer referencia a fuentes adicionales de datos y cuantificar predicados de comparación. El capítulo describe los diferentes tipos de predicados y muestra cómo son utilizados para recuperar datos específicos de una base de datos SQL.

Capítulo 10: Trabajar con funciones y expresiones de valor

Este capítulo explica cómo utilizar los diferentes tipos de funciones y expresiones de valor en las instrucciones SQL. Se aprenderá cómo utilizar funciones set, funciones de valor, expresiones de valor y valores especiales en diferentes cláusulas dentro de la instrucción SQL.

Capítulo 11: Acceder a múltiples tablas

Este capítulo describe cómo unir tablas para poder recuperar datos desde esas tablas. Se aprenderá cómo realizar operaciones join básicas, unir tablas con nombres de columna compartidos, utilizar la función join de condición y realizar operaciones de unión.

Capítulo 12: Utilizar subconsultas para acceder y modificar datos

En este capítulo se aprenderá cómo crear subconsultas que arrojen múltiples filas y otras que arrojen solamente un valor. También se aprenderá cómo utilizar subconsultas correlacionadas y subconsultas anidadas. Adicionalmente, se aprenderá cómo utilizar subconsultas para modificar datos.

PARTE III Acceso avanzado a los datos

La parte III presenta una introducción a las técnicas avanzadas de acceso de datos como las rutinas invocadas por SQL, activadores y cursores. También se aprenderá cómo administrar transacciones, cómo acceder a los datos SQL desde el programa host y cómo incorporar datos XML a la base de datos.

Capítulo 13: Crear rutinas invocadas por SQL

Este capítulo describe los procedimientos y las funciones invocadas por SQL y cómo pueden crearse en la base de datos SQL. Se aprenderá cómo definir parámetros de entrada, agregar variables locales a la rutina, trabajar con instrucciones de control y utilizar parámetros de salida.

Capítulo 14: Crear activadores SQL

Este capítulo presenta los activadores SQL y explica cómo crear activadores de inserción, de actualización y de eliminación en la base de datos SQL. Se aprenderá cómo los activadores son invocados automáticamente y qué tipos de acciones se pueden tomar.

Capítulo 15: Utilizar cursores SQL

En este capítulo se aprenderá cómo los cursores SQL se utilizan para recuperar una fila de datos a la vez desde un conjunto de resultados. El capítulo explica cómo declarar un cursor, abrir y cerrar un cursor, y cómo recuperar datos desde un cursor. También se aprenderá cómo utilizar instrucciones UPDATE y DELETE posicionadas después de buscar una fila utilizando un cursor.

Capítulo 16: Manejar transacciones SQL

En este capítulo se aprenderá cómo se utilizan las transacciones para asegurar la integridad de los datos SQL. El capítulo describe cómo establecer las propiedades de la transacción, iniciar una transacción, establecer el desplazamiento de las restricciones, crear puntos de recuperación en una transacción y finalizar una transacción.

Capítulo 17: Acceder a datos SQL desde un programa host

Este capítulo describe los cuatro métodos soportados por el estándar SQL para acceder a una base de datos SQL. Se aprenderá cómo invocar SQL directamente desde una aplicación cliente, incrustar instrucciones SQL en un programa, crear módulos cliente de SQL y utilizar una interfaz de nivel de llamada de SQL para acceder a los datos.

Capítulo 18: Trabajar con datos XML

Este capítulo describe cómo los datos XML pueden incorporarse a una base SQL. Aprenderá los aspectos básicos de XML, cómo usar los tipos de datos de XML para almacenar XML en los valores de columna de la tabla, escribir funciones SQL/XML para arrojar datos desde las bases formateadas como XML, y las reglas de trazado de SQL/XML que describen cómo los valores SQL son transformados a XML y viceversa.

PARTE IV Apéndices**Apéndice A: Respuestas a los autoexámenes**

Este apéndice proporciona las respuestas para las preguntas de las autoevaluaciones enlistadas al final de cada capítulo.

Apéndice B: Palabras clave de SQL:2006

Este apéndice enumera las palabras clave reservadas y no reservadas de la forma en que se utilizan en las instrucciones SQL, como se define en el estándar SQL:2006.

Apéndice C: Código SQL utilizado en los ejercicios Pruebe esto

Este apéndice enumera todo el código SQL utilizado en los ejercicios de este libro, consolidados en un solo lugar para una referencia más fácil. Este código también puede ser descargado desde www.mcgraw-hill-educacion.com. Haga una búsqueda por autor, ISBN o título.

Contenido de cada capítulo

Como se puede observar en el resumen, *Fundamentos de SQL* está organizado en capítulos. Cada capítulo se enfoca en un conjunto de tareas relacionadas. El capítulo contiene la información de importancia que se necesita para comprender los diferentes conceptos relacionados con esas tareas, explica cómo crear las instrucciones SQL necesarias para realizar las tareas y proporciona ejemplos de cómo se crean esas instrucciones. Suplementariamente, cada capítulo contiene elementos adicionales para ayudar a comprender mejor la información cubierta en ese capítulo:

- **Pregunte al experto** Cada capítulo contiene una o dos de estas secciones; proporcionan información acerca de preguntas que pudieran surgir referentes a la información presentada en el capítulo.
- **Autoexamen** Cada capítulo termina con un Autoexamen, que es un conjunto de preguntas que examinan al usuario acerca de la información y habilidades que se aprendieron en este capítulo. Las respuestas a los autoexámenes se encuentran en el apéndice A.

Sintaxis SQL

La sintaxis de una instrucción SQL se refiere a la estructura y a las reglas utilizadas para esa instrucción, como se especifica en SQL:2006. La mayoría de los capítulos incluirán la sintaxis para una o más instrucciones para que se pueda tener un entendimiento de los elementos básicos contenidos en esas instrucciones. Por ejemplo, la siguiente sintaxis representa la información necesaria cuando se define una instrucción CREATE TABLE:

```
<definición de la tabla> ::=  
CREATE [ { GLOBAL | LOCAL } TEMPORARY ] TABLE <nombre de la tabla>  
(<elemento de la tabla> [ { , <elemento de la tabla> } ... ] )  
[ ON COMMIT { PRESERVE | DELETE } ROWS ]
```

NOTA

Por el momento no tiene que preocuparse acerca del significado de este código SQL. Se presenta este ejemplo solamente para mostrar cómo se representan las instrucciones SQL en este libro.

Como se puede ver, la sintaxis de una instrucción puede contener muchos elementos. Nótese que la mayoría de las palabras utilizadas dentro de la instrucción se muestran en mayúsculas. Las palabras en mayúsculas son palabras clave SQL que se utilizan para formular la instrucción SQL. (Para una lista completa de las palabras clave SQL:2006, vea el apéndice B.) A pesar de que SQL no requiere que las palabras clave sean escritas en mayúsculas, se utiliza esa convención en este libro para que pueda identificar fácilmente las palabras clave dentro de una instrucción. Adicionalmente a las palabras clave, la sintaxis para una instrucción SQL incluye otros muchos elementos que ayudan a definir cómo deberá ser creada una instrucción en particular:

- **Corchetes** Los corchetes indican que la sintaxis encerrada en esos corchetes es opcional. Por ejemplo, la cláusula ON COMMIT en la instrucción CREATE TABLE es opcional.
- **Corchetes angulares** Los corchetes angulares encierran información que representa un marcador de posición. Cuando la instrucción ha sido creada, el marcador de posición es reemplazado por los

elementos o identificadores SQL apropiados. Por ejemplo, deberá reemplazarse el marcador de posición <nombre de la tabla> con un nombre para la tabla cuando se defina la instrucción CREATE TABLE.

- **Llaves** Las llaves se utilizan para agrupar elementos. Las llaves indican que primero se deberá decidir cómo manejar los contenidos dentro de las llaves y luego determinar cómo encajan dentro de la instrucción. Por ejemplo, el conjunto de palabras clave PRESERVE | DELETE está encerrado entre llaves. Primero se deberá elegir entre PRESERVE o DELETE y luego tratar con la línea completa del código. Como resultado, la cláusula puede leerse como ON COMMIT PRESERVE ROWS, o como ON COMMIT DELETE ROWS.
- **Barras verticales** La | Puede leerse como “o”, lo que significa que deberá utilizar ya sea la opción PRESERVE o la opción DELETE.
- **Tres puntos** Los tres puntos indican que se puede repetir la cláusula tantas veces como sea necesario. Por ejemplo, se pueden incluir tantos elementos de la tabla como sean necesarios (representados por <elemento de la tabla>).
- **Dos puntos y signo de igual** El símbolo ::= (dos veces consecutivas los dos puntos y el signo de igual) indica que el marcador de posición a la izquierda del símbolo está definido por la sintaxis que sigue al símbolo. En el ejemplo de la sintaxis, el marcador de posición <definición de la tabla> iguala a la sintaxis que conforma una instrucción CREATE TABLE.

Al referirse a la sintaxis deberá ser capaz de construir una instrucción SQL que cree objetos de la base de datos o modifique datos SQL según sea necesario. Sin embargo, para poder demostrar mejor cómo se aplica la sintaxis, cada capítulo contiene también ejemplos de instrucciones SQL reales.

Ejemplos de instrucciones SQL

Cada capítulo proporciona ejemplos de cómo se implementan las instrucciones SQL cuando se accede a una base de datos SQL. Por ejemplo, podría observarse una instrucción SQL similar a la siguiente:

```
CREATE TABLE ARTISTAS
( ID_ARTISTA          INT,
  NOMBRE_ARTISTA      VARCHAR(60) ,
  FDN_ARTISTA         DATE,
  POSTER_EN_EXISTENCIA  BOOLEAN );
```

Observe que la instrucción está escrita en un tipo especial para mostrar que se trata de código SQL. Observe también que las palabras clave y los nombres de objeto están todos en mayúsculas. (Hasta este punto, no es necesario preocuparse por ninguno de los demás detalles.)

Los ejemplos utilizados en este libro son de SQL puro, lo que significa que están basados en el estándar SQL:2006. Se encontrará, sin embargo, que en algunos casos las implementaciones SQL no soportan una instrucción SQL exactamente de la misma forma en la que está definida en el estándar. Por esta razón, también pudiera ser necesario referirse a la documentación para un producto en particular con tal de asegurarse de que sus instrucciones SQL acaten la implementación de ese producto de SQL.

Pudiera ser a veces solamente una variación ligera, pero habrá ocasiones en las que la instrucción del producto será sustancialmente diferente de la instrucción SQL estándar.

Los ejemplos en cada capítulo están basados en una base de datos relacionada a un inventario de discos compactos. Sin embargo, los ejemplos no son necesariamente consistentes en términos de los nombres utilizados para los objetos de la base de datos y cómo esos objetos están definidos. Por ejemplo, dos capítulos diferentes pudieran contener ejemplos que hacen referencia a una tabla llamada INVENTARIO_CD. Sin embargo, no se puede asumir que las tablas utilizadas en diferentes ejemplos estén conformadas por las mismas columnas o contengan los mismos datos. Debido a que cada ejemplo se enfoca en un aspecto único de SQL, las tablas utilizadas en los ejemplos están definidas en una forma específica para las necesidades de ese ejemplo, como se verá mientras se avanza en los capítulos. Éste no es el caso para los ejercicios, que utilizan una estructura de base de datos consistente a través de todo el libro.

Ejercicios Pruebe esto

Cada capítulo contiene uno o dos ejercicios que permiten aplicar la información que se aprendió durante el capítulo. Cada ejercicio está dividido en pasos que lo llevarán a través del proceso de completar una tarea en particular. Muchos de los proyectos incluyen archivos relacionados que pueden ser descargados desde nuestro sitio Web en www.mcgraw-hill-educacion.com. Los archivos a menudo incluyen las instrucciones SQL utilizadas dentro de los ejercicios Pruebe esto (**en inglés**). Adicionalmente, se incluye una consolidación (**en español**) de estas instrucciones SQL en el apéndice C.

Los ejercicios están basados en la base de datos INVENTARIO. Usted creará la base de datos y las tablas en otros objetos en la base de datos, agregará datos a esas tablas, y luego manipulará esos datos. Debido a que los proyectos se construyen uno sobre el otro, es mejor que los complete en el orden en que están presentados en el libro. Esto resulta de especial importancia para los capítulos en la parte I, en la cual se crean los objetos de la base de datos, y en el capítulo 7, en el cual se insertan los datos en las tablas. Sin embargo, si de todas maneras usted prefiere hacerlos de forma aleatoria, puede referirse al apéndice C, que proporciona todo el código necesario para crear los objetos de la base de datos y poblar esas tablas con datos.

Para completar la mayoría de los ejercicios de este libro, usted necesitará tener acceso a un RDBMS que le permita ingresar y ejecutar instrucciones SQL interactivamente. Si está accediendo a un RDBMS a través de una red, verifique que el administrador de la base de datos se asegure de que usted está ingresando con las credenciales necesarias para crear una base de datos y un esquema. Podría necesitar permisos especiales para crear esos objetos. También verifique si es que existen ciertos parámetros que deberá incluir cuando cree la base de datos (por ejemplo, el tamaño de archivo de ingreso), restricciones acerca de los nombres que puedan utilizarse o restricciones de cualquier otro tipo. Asegúrese de revisar la documentación del producto antes de trabajar con cualquier producto de base de datos.

Parte I

Bases de datos relacionales y SQL

Capítulo 1

Introducción a las
bases de datos
relacionales y a SQL

Habilidades y conceptos clave

- Entienda las bases de datos relacionales
 - Aprenda acerca de SQL
 - Use un sistema de gestión de base de datos relacional
-

En el 2006, la International Organization for Standardization (ISO) y el American National Standards Institute (ANSI) publicaron revisiones a su estándar SQL, al cual llamaré SQL:2006. Como verá después, el estándar está dividido en partes, y cada parte es aprobada y publicada en su propia línea de tiempo; por lo tanto, diferentes partes tienen distintos años de publicación. Es común usar el último año como nombre colectivo para el conjunto de todas las partes publicadas durante ese año. El estándar SQL 2006, como sus predecesores SQL:2003, SQL:1999 (también conocido como SQL3) y SQL-92, se basa en el modelo de datos relacional, el cual define cómo los datos pueden ser almacenados y manipulados dentro de una base de datos relacional. Los sistemas de gestión de base de datos relacional (RDBMS) como Oracle, Sybase, DB2, MySQL y Microsoft SQL Server (o sólo SQL Server) usan el estándar SQL como base de su tecnología, proporcionando entornos de base de datos que apoyan tanto a SQL como al modelo de datos relacional. Hay más información sobre el estándar SQL más adelante en este capítulo.

Entienda las bases de datos relacionales

El lenguaje estructurado de consultas (SQL, Structured Query Language) apoya la creación y mantenimiento de la base de datos relacional y la gestión de los datos dentro de la base de datos. Sin embargo, antes de entrar en discusión acerca de las bases de datos relacionales, quiero explicar el significado del término *base de datos*. El término ha sido utilizado para referirse a cualquier cosa, desde una colección de nombres y direcciones hasta un complejo sistema de recuperación y almacenamiento de datos que se basa en interfaces de usuarios y una red de computadoras y servidores. Hay tantas definiciones para la palabra *base de datos* como libros sobre éstas. Por otra parte, los distintos proveedores de DBMS han desarrollado diferentes arquitecturas, por lo que no todas las bases de datos están diseñadas de la misma manera. A pesar de la falta de una definición absoluta, la mayoría de las fuentes coinciden en que una base de datos, por lo menos, es una colección de datos organizada en un formato estructurado que es definido como *metadatos* que describe esa estructura. Puede pensar en los metadatos como información sobre los datos almacenados, que define cómo se almacenan éstos en una base de datos.

A lo largo de los años se ha implementado una serie de modelos de base de datos para almacenar y administrar la información. Varios de los modelos más comunes incluyen los siguientes:

- **Jerárquico** Este modelo tiene una estructura primario y secundario que es similar a un árbol invertido, que es lo que constituye la jerarquía. Los datos están organizados en *nodos*, el equivalente lógico de tablas en una base de datos relacional. Un nodo principal puede tener muchos nodos secundarios, pero un nodo secundario sólo puede tener un nodo principal. Aunque

el modelo ha sido muy utilizado, a menudo se considera inadecuado para muchas aplicaciones debido a su estructura inflexible y la falta de apoyo de relaciones complejas. Aun algunas implementaciones como IMS de IBM han introducido características que trabajan en torno a estas limitaciones.

- **Red** Este modelo aborda algunas de las limitaciones del modelo jerárquico. Los datos están organizados en *tipos de registro*, el equivalente lógico de tablas en una base de datos relacional. Al igual que el modelo jerárquico, el modelo de red usa la estructura de un árbol invertido, pero los tipos de registro se organizan en una estructura que relaciona pares de tipos de registro en propietarios y miembros. Cualquier tipo de registro puede participar en cualquier conjunto con otros tipos de registro en la base de datos, que apoya a las consultas y relaciones más complejas de lo que es posible en el modelo jerárquico. Hasta el modelo de red tiene sus limitaciones, y la más seria es la complejidad. Al acceder a la base de datos, el usuario debe estar familiarizado con la estructura y mantener un seguimiento cuidadoso de dónde está y cómo llegó ahí. También es difícil cambiar la estructura sin afectar las aplicaciones que interactúan con la base de datos.
- **Relacional** Este modelo aborda algunas de las limitaciones de los modelos jerárquicos y de red. En una base de datos de modelo jerárquico o de red, la aplicación se basa en una implementación determinada de esa base de datos, que luego es codificada en la aplicación. Si agrega un nuevo atributo (elemento de datos) a la base de datos, debe modificar la aplicación, aun cuando no se use ese atributo. Sin embargo, una base de datos relacional es independiente de la aplicación; puede hacer modificaciones no destructivas a la estructura sin afectar la aplicación. Además, la estructura de la base de datos relacional se basa en la relación, o tabla, junto con la habilidad de definir relaciones complejas entre ellas. Se puede acceder directamente a cada relación sin la lentitud de las limitaciones de los modelos jerárquicos o propietario/miembro que requiere de una navegación a través de una estructura compleja de datos. En la siguiente sección, “El modelo relacional”, se verá con mayor detalle este modelo.

Aunque aún se usan en muchas organizaciones, las bases de datos de modelo jerárquico y de red ahora se consideran como soluciones heredadas. El modelo relacional es el más ampliamente aplicado en los sistemas de negocios modernos, y es el modelo relacional el que proporciona la base para SQL.

El modelo relacional

Si alguna vez tiene la oportunidad de ver un libro acerca de base de datos relacionales, es muy posible que vea el nombre de E. F. (Ted) Codd, a quien se hace referencia en el contexto del modelo relacional. En 1970, Codd publicó su trabajo más importante “A Relational Model Of Data For Large Shared Data Banks” (Un modelo relacional de datos para grandes bancos de datos compartidos), en el diario *Communications of the ACM*, volumen 13, número 6 (junio de 1970). Codd define una estructura de datos relacional que protege los datos y permite que sean manipulados de manera que es previsible y resistente al error. El modelo relacional, el cual se basa principalmente en los principios matemáticos de la teoría de conjuntos y lógica de predicados, apoya la recuperación de datos sencilla, aplica la *integración de datos* (la precisión y coherencia de los datos), y proporciona una estructura de base de datos independiente de las aplicaciones al acceder a los datos almacenados.

El núcleo del modelo relacional es la relación. Una *relación* es un conjunto de columnas y filas reunidas en una estructura en forma de tabla que representa una entidad única formada por los datos relacionados. Una *entidad* es una persona, lugar, cosa, evento o concepto sobre el cual los datos son recolectados, como un artista, un libro o una transacción de ventas. Cada relación comprende uno o más atributos (columnas). Un *atributo* es un hecho simple que describe o caracteriza una entidad de alguna manera. Por ejemplo, en la figura 1-1, la entidad es un disco compacto (CD) con atributos de NOMBRE_CD (el título del CD), NOMBRE_ARTISTA (el nombre del artista) y AÑO_DERECHOSDEAUTOR (el año de los derechos de autor del disco).

Como se ve en la figura 1-1, cada atributo tiene un dominio asociado. Un *dominio* define el tipo de datos que son almacenados en un atributo particular; sin embargo, un dominio no es lo mismo que un tipo de datos. Un *tipo de datos*, el cual se tratará con mayor detalle en el capítulo 3, es un tipo específico de *restricción* (un control usado para hacer cumplir la integridad de los datos) asociados con una columna, mientras que un dominio, tal como se utiliza en el modelo relacional, tiene un significado más amplio y describe exactamente qué datos pueden ser incluidos en un atributo asociado con ese dominio. Por ejemplo, el atributo AÑO_DERECHOSDEAUTOR es asociado con el dominio de Año. Como se ve en este ejemplo, es común la práctica de incluir una clase de palabra que describe los nombres de dominio en el atributo, pero esto no es obligatorio. El dominio puede ser definido de manera que el atributo sólo incluya datos cuyos valores y formato sean limitados a años, a diferencia de días o meses. El dominio también puede limitar los datos a un rango específico de años. Un tipo de datos, por otro lado, restringe el formato de los datos, como el permitir únicamente dígitos numéricos, pero no los valores, a menos que esos valores violen de alguna manera el formato.

Los datos se almacenan en una relación en tuplas (filas). Una *tupla* es un conjunto de datos cuyos valores hacen una instancia de cada atributo definido por esa relación. Cada tupla representa un registro de datos relacionados. (De hecho, el conjunto de datos se conoce en ocasiones como *registro*.) Por ejemplo, en la figura 1-1 la segunda tupla de arriba hacia abajo contiene el valor “Joni Mitchell” para el atributo NOMBRE_ARTISTA, el valor “Blue” para el atributo NOMBRE_CD y el valor “1971” para el atributo AÑO_DERECHOSDEAUTOR. Estos tres valores juntos forman una tupla.

Atributo	Nombre del atributo (NOMBRE_CD)	Nombre del dominio (Año)
NOMBRE_ARTISTA:NombreCompleto	NOMBRE_CD:Título	AÑO_DERECHOSDEAUTOR:Año
Jennifer Warnes	Famous Blue Raincoat	1991
Joni Mitchell	Blue	1971
William Ackerman	Past Light	1983
Kitaro	Kojiki	1990
Bing Crosby	That Christmas Feeling	1993
Patsy Cline	Patsy Cline: 12 Greatest Hits	1988

Figura 1-1 Relación con los atributos NOMBRE_CD, NOMBRE_ARTISTA y AÑO_DERECHOSDEAUTOR.

NOTA

Los términos lógicos relación, atributo y tupla se usan principalmente para referirse al modelo relacional. SQL usa los términos físicos tabla, columna y fila para describir esos elementos. Debido a que el modelo relacional se basa en modelos matemáticos (un modelo lógico) y SQL se enfoca más a la implementación física del modelo, los significados para los términos del modelo y los términos de lenguaje de SQL son ligeramente distintos, pero los principios básicos son los mismos. Los términos de SQL se analizan con más detalle en el capítulo 2.

El modelo relacional es, por supuesto, más complejo que los meros atributos y tuplas que hacen la relación. Dos consideraciones importantes en el diseño e implementación de cualquier base de datos relacional son la normalización de los datos y las asociaciones de relaciones entre los distintos tipos de datos.

Normalización de datos

La parte central de los principios del modelo relacional es el concepto de *normalización*, una técnica para producir un conjunto de relaciones que poseen un conjunto de ciertas propiedades que minimizan los datos redundantes y preservan la integridad de los datos almacenados tal como se mantienen (añadidos, actualizados y eliminados). El proceso fue desarrollado por E. F. Codd en 1972, y el nombre es un chiste político debido a que el presidente Nixon estaba “normalizando” relaciones con China en ese momento. Codd imaginó que si las relaciones con un país pueden normalizarse, entonces seguramente podría normalizar las relaciones de la base de datos. La normalización se define por un conjunto de normas, que se conocen como *formas normales*, que proporcionan una directriz específica de cómo los datos son organizados para evitar anomalías que den lugar a inconsistencias y pérdida de los datos tal como se mantienen almacenados en la base de datos.

Cuando Codd presentó por primera vez la normalización, incluía tres formas normales. A pesar de que formas normales adicionales se han agregado desde entonces, las tres primeras cubren la mayoría de las situaciones que se encontrarán en las dos bases de datos personales y empresariales, y ya que la intención principal es presentar el proceso de normalización, sólo se analizarán esas tres formas.

Elección de un identificador único Un identificador único es un atributo o conjunto de atributos que únicamente identifican cada fila de datos en una relación. El identificador único eventualmente se convertirá en la clave principal de la tabla creada en la base de datos física desde la relación de normalización, pero muchos usan los términos *identificador único* y *clave principal* de manera intercambiable. A cada identificador potencial único se le denomina *candidato clave*, y cuando hay varios candidatos, el diseñador elegirá el mejor, el cual es el menos probable de cambiar valores o el más simple y/o el más corto. En muchos casos, un solo atributo se identifica únicamente en los datos en cada tupla de la relación. Sin embargo, cuando ningún atributo que es único se encuentra, el diseñador busca varios atributos que puedan ser unidos (puestos juntos) para formar un identificador único. En los pocos casos donde ningún candidato clave razonable es encontrado, el diseñador debe inventar un identificador único denominado *sustituto clave*, frecuentemente con valores asignados al azar o secuencialmente cuando las tuplas sean agregadas a la relación.

Mientras no sea absolutamente necesario hasta la segunda forma normal, es habitual seleccionar un identificador único como primer paso en la normalización. Es más fácil de esa manera.

Primera forma normal La primera forma normal, que proporciona la fundación para la segunda y tercera forma normal, incluye las siguientes directrices:

- Cada atributo de una tupla contiene sólo un valor.
- Cada tupla en una relación contiene el mismo número de atributos.
- Cada tupla es diferente, lo que significa que la combinación de los valores de todos los atributos de una tupla dada no puede ser como ninguna otra tupla en la misma relación.

Como se ve en la figura 1-2, la segunda tupla y la última tupla violan la primera forma normal. En la segunda tupla, el atributo NOMBRE_CD y el atributo AÑO_DERECHOSDEAUTOR contienen cada una dos valores. En la última tupla, el atributo NOMBRE_ARTISTA contiene tres valores. También debe vigilar la repetición de valores en forma de repetición de columnas. Por ejemplo, dividir el atributo NOMBRE_ARTISTA en tres atributos llamados NOMBRE_ARTISTA_1, NOMBRE_ARTISTA_2 y NOMBRE_ARTISTA_3 no es una solución adecuada porque es muy probable que necesite un cuarto nombre, luego un quinto, y así sucesivamente. Además, repetir columnas hace que las consultas sean más difíciles, ya que debe recordar buscar en todas las columnas cuando busque un valor específico.

Para normalizar la relación mostrada en la figura 1-2, debe crear relaciones adicionales que separen los datos de modo que cada atributo contenga un solo valor, cada tupla contenga el mismo número de atributos y cada tupla sea diferente, como se muestra en la figura 1-3. Ahora los datos se ajustan a la primera forma normal.

Observe que hay valores duplicados en la segunda relación; el valor de 10002 en el ID_ARTISTA se repite y el valor de 99308 en el ID_CD también se repite. Sin embargo, cuando el valor de los dos atributos en cada tupla son puestos juntos, la tupla como un todo forma una combinación única, lo cual significa que, a pesar de la aparente duplicidad, cada tupla en la relación es diferente.

NOMBRE_ARTISTA	NOMBRE_CD	AÑO_DERECHOSDEAUTOR
Jennifer Warnes	Famous Blue Raincoat	1991
Joni Mitchell	Blue; Court and Spark	1971; 1974
William Ackerman	Past Light	1983
Kitaro	Kojiki	1990
Bing Crosby	That Christmas Feeling	1993
Patsy Cline	Patsy Cline: 12 Greatest Hits	1988
Jose Carreras; Placido Domingo; Luciano Pavarotti	Carreras Domingo Pavarotti in Concert	1990

Figura 1-2 Relación de la violación de la primera forma normal.

ID_ARTISTA	NOMBRE_ARTISTA	ID_ARTISTA	ID_CD	ID_CD	NOMBRE_CD	AÑO_DERECHOSDEAUTOR
10001	Jennifer Warnes	10001	99301	99301	Famous Blue Raincoat	1991
10002	Joni Mitchell	10002	99302	99302	Blue	1971
10003	William Ackerman	10002	99303	99303	Court and Spark	1974
10004	Kitaro	10003	99304	99304	Past Light	1983
10005	Bing Crosby	10004	99305	99305	Kojiki	1990
10006	Patsy Cline	10005	99306	99306	That Christmas Feeling	1993
10007	Jose Carreras	10006	99307	99307	Patsy Cline: 12 Greatest Hits	1988
10008	Plácido Domingo	10007	99308	99308	Carreras Domingo Pavarotti in Concert	1990
10009	Luciano Pavarotti	10008	99308			
		10009	99308			

Figura 1-3 Relación que se ajusta a la primera forma normal.

¿Observó que los atributos ID_ARTISTA y ID_CD fueron agregados? Se hizo esto porque no había otros candidatos clave. NOMBRE_ARTISTA no es único (dos personas con el mismo nombre pueden ser artistas), y tampoco es NOMBRE_CD (dos CD pueden terminar con el mismo nombre, aunque es probable que sean de diferentes sellos discográficos). ID_ARTISTA es la clave principal de la primera relación, e ID_CD es la clave principal en la tercera. La clave principal en la segunda relación es la combinación de ID_ARTISTA y ID_CD.

Segunda forma normal Para comprender la segunda forma normal, primero debe entender el concepto de *dependencia funcional*. Para esta definición se usarán dos atributos arbitrarios, hábilmente llamados A y B. El atributo B es *funcionalmente dependiente* (*dependiente* para abreviar) del atributo A si en cualquier momento no hay más que un valor del atributo B asociado con el valor dado al atributo A. Para que no te preguntes de qué planeta vengo, tratemos de hacer la definición más entendible. Si se dice que el atributo B es funcionalmente dependiente del atributo A, también estaremos diciendo que el atributo A *determina* al atributo B, o que A es un factor *determinante* (identificador único) del atributo B. En la figura 1-4, AÑO_DERECHOSDEAUTOR depende de ID_CD, ya que sólo puede haber un valor de AÑO_DERECHOSDEAUTOR para cualquier CD. Dicho de otra manera, ID_CD es un factor determinante de AÑO_DERECHOSDEAUTOR.

La segunda forma normal expone que una relación debe estar en la primera forma normal y que todos los atributos en la relación dependen del identificador único *completo*. En la figura 1-4, si la combinación de ID_ARTISTA y ID_CD es seleccionada como identificador único, entonces AÑO_DERECHOSDEAUTOR violaría la segunda forma normal porque sólo dependería de ID_CD en lugar de la combinación ID_CD y ID_ARTISTA. A pesar de que la relación se ajusta a la primera forma normal, se violaría la segunda forma normal. De nuevo, la solución sería separar los datos en relaciones diferentes, como se vio en la figura 1-3.

Tercera forma normal La tercera forma normal, como la segunda forma normal, depende de la relación del identificador único. Para adherir a las directrices de la tercera forma normal, una

← Identificador único →		
ID_ARTISTA	ID_CD	AÑO_DERECHOSDEAUTOR
10001	99301	1991
10002	99302	1971
10002	99303	1974
10003	99304	1983
10004	99305	1990
10005	99306	1993
10006	99307	1988

Figura 1-4 Relación con un identificador único concatenado.

relación debe estar en la segunda forma normal y sin un atributo clave (atributos que no sean parte de algún candidato clave) deben ser independiente el uno del otro y depender del identificador único. Por ejemplo, el identificador único en la relación mostrada en la figura 1-5 es el atributo ID_ARTISTA.

Identificador único			
ID_ARTISTA	NOMBRE_ARTISTA	ID_AGENCIA	ESTADO_AGENCIA
10001	Jennifer Warnes	2305	NY
10002	Joni Mitchell	2306	CA
10003	William Ackerman	2306	CA
10004	Kitaro	2345	NY
10005	Bing Crosby	2367	VT
10006	Patsy Cline	2049	TN
10007	Jose Carreras	2876	CA
10008	Placido Domingo	2305	NY
10009	Luciano Pavarotti	2345	NY

Figura 1-5 Relación con un atributo que viola la tercera forma normal.

Los atributos NOMBRE_ARTISTA e ID_AGENCIA dependen del identificador único y son independientes uno del otro. Sin embargo, el atributo ESTADO_AGENCIA depende del atributo ID_AGENCIA, y por lo tanto viola las condiciones de la tercera forma normal. Este atributo se adapta mejor en una relación que incluye datos sobre las agencias.

NOTA

En el mundo teórico del diseño relacional, el objetivo es almacenar los datos de acuerdo con las reglas de normalización. Sin embargo, en el mundo real de aplicación de bases de datos, ocasionalmente se desnormalizan los datos, lo que significa que se violen deliberadamente las reglas de normalización, particularmente la segunda y la tercera forma normal. La desnormalización se usa principalmente para mejorar el rendimiento o reducir la complejidad en los casos en donde una estructura demasiado normalizada complica la implementación. Con todo, el objetivo de la normalización es asegurar la integridad de los datos; por lo tanto, la desnormalización se realizaría con sumo cuidado y como último recurso.

Relaciones

Hasta ahora el enfoque en este capítulo se ha centrado en la relación y la manera de normalizar los datos. Sin embargo, un componente importante de cualquier base de datos relacional es de qué forma esas relaciones se asocian entre sí. Esas asociaciones, o *relaciones*, se vinculan en forma significativa, lo que contribuye a garantizar la integridad de los datos de modo que una acción realizada en una relación no repercuta negativamente en los datos de otra relación.

Hay tres tipos principales de relaciones:

- **Una a una** Una relación entre dos relaciones en la cual una tupla en la primera relación esté relacionada con al menos una tupla en la segunda relación, y una tupla en la segunda relación esté relacionada con al menos una tupla en la primera relación.
- **Una a varias** Una relación entre dos relaciones en la cual una tupla en la primera relación esté relacionada con ninguna, una o más tuplas en la segunda relación, pero una tupla en la segunda relación esté relacionada con al menos una tupla en la primera relación.
- **Varias a varias** Una relación entre dos relaciones en la cual una tupla en la primera relación esté relacionada con ninguna, una o más tuplas en la segunda relación, y una tupla en la segunda relación esté relacionada con ninguna, una o más tuplas en la primera relación.

La mejor manera de ilustrar estas relaciones es ver un modelo de datos de varias relaciones (mostrado en la figura 1-6). Las relaciones se nombraron para hacer referencia a ellas con mayor facilidad.

- Una relación una a una existe entre las relaciones AGENCIAS_ARTISTA y NOMBRES_ARTISTA. Para cada uno de los artistas que se listan en la relación AGENCIAS_ARTISTA, sólo puede haber una tupla correspondiente en la relación NOMBRES_ARTISTA, y viceversa. Esto implica una regla de negocio que un artista puede trabajar con sólo una agencia a la vez.

- Una relación una a varias existe entre las relaciones NOMBRE_ARTISTA y CD_ARTISTA. Para cada uno de los artistas en la relación NOMBRES_ARTISTA, ninguna, una o más tuplas para ese artista pueden estar en la lista en la relación CD_ARTISTA. En otras palabras, cada artista pudo haber hecho ninguno, uno o más CD. Sin embargo, para cada uno de los artistas que figuran en la relación CD_ARTISTA, sólo puede haber una tupla correspondiente para cada artista en la relación NOMBRES_ARTISTA debido a que cada artista puede tener sólo una tupla en la relación NOMBRES_ARTISTA.
- Una relación una a varias existe entre las relaciones CD_ARTISTA y DISCOS_COMPACTOS. Para cada CD, puede haber uno o más artistas; sin embargo, cada tupla en CD_ARTISTA puede corresponder en sólo una tupla en DISCOS_COMPACTOS debido a que cada CD sólo aparece una vez en la relación DISCOS_COMPACTOS.
- Una relación varias a varias existe entre las relaciones NOMBRE_ARTISTA y DISCOS_COMPACTOS. Para cada artista, puede haber ninguno, uno o más CD, y para cada CD, puede haber uno o más artistas.

NOTA

Las bases de datos relacionales sólo apoyan una relación una a varias directamente. Una relación varias a varias se implementa físicamente agregando una tercera relación entre la primera y la segunda para crear dos relaciones una a varias. En la figura 1-6, la relación CD_ARTISTA se agregó entre las relaciones NOMBRES_ARTISTA y DISCOS_COMPACTOS. Una relación una a una se aplica físicamente al igual que una relación una a varias, excepto que se añade una limitación para evitar duplicar los registros que coinciden en los “muchos” lados de la relación. En la figura 1-6 se añadió una limitación única en el atributo ID_ARTISTA para evitar que un artista aparezca en más de una agencia.

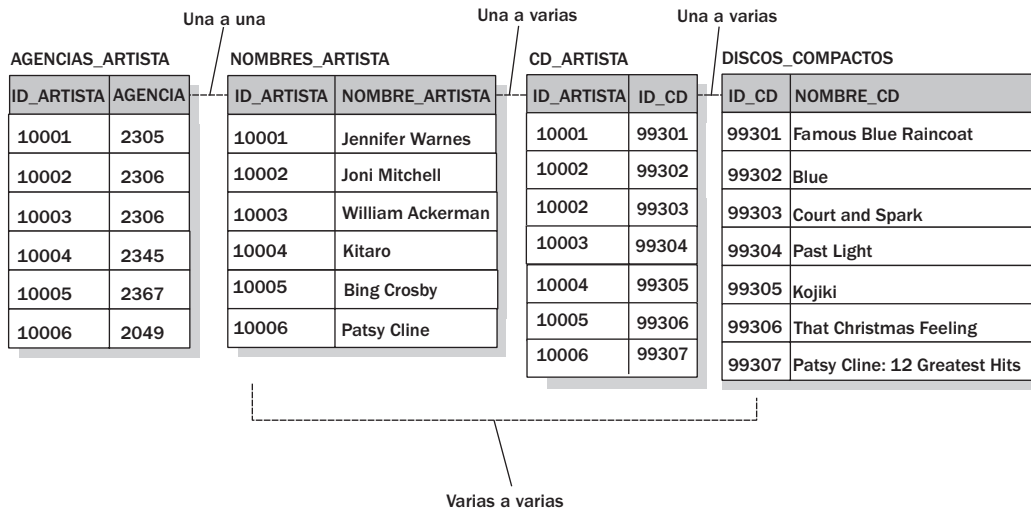


Figura 1-6 Tipos de relaciones entre relaciones.

Las relaciones se clasifican también por *cardinalidad mínima* (el número mínimo de tuplas que participan en una relación). Si cada tupla en una relación *debe* tener una tupla correspondiente en la otra, la relación se dice que es *obligatoria* en esa dirección. Del mismo modo, si cada tupla en una relación no requiere una tupla correspondiente en la otra, la relación se dice que es *opcional* en esa dirección. Por ejemplo, la relación entre NOMBRES_ARTISTA y AGENCIAS_ARTISTA es obligatoria-obligatoria porque cada artista debe tener una agencia y cada tupla de AGENCIAS_ARTISTA debe referirse a un único artista. Las reglas de negocio deben entenderse antes que la cardinalidad mínima se determine con certeza. Por ejemplo, ¿podemos tener un artista en la base de datos que en algún momento no tenga CD en la base de datos (es decir, ninguna tupla correspondiente en CD_ARTISTA)? Si es así, entonces la relación entre NOMBRES_ARTISTA y CD_ARTISTA es obligatoria-opcional; de lo contrario, es obligatoria-obligatoria.

Pregunta al experto

P: Se mencionó que las relaciones entre relaciones contribuyen a garantizar la integridad de los datos. ¿Cómo pueden las relaciones hacerlo posible?

R: Suponga que su modelo de datos incluye una relación (llamada NOMBRE_ARTISTA) que muestre en su inventario todos los artistas que han grabado CD. Su modelo también incluye una relación (CD_ARTISTA) donde coinciden los ID de los artistas con los ID de los discos compactos. Si existe una relación entre las dos relaciones, las tuplas de una relación siempre corresponderán a las tuplas de la otra relación. Como resultado, se podrán evitar ciertas acciones que puedan comprometer los datos. Por ejemplo, no sería posible añadir el ID de un artista a la relación CD_ARTISTA si ese ID no se muestra en la lista en la relación NOMBRE_ARTISTA. Tampoco sería posible borrar un artista de la relación NOMBRE_ARTISTA si el ID del artista hace referencia a la relación CD_ARTISTA.

P: ¿A qué se refiere con el término *modelo de datos*?

R: Con modelo de datos me refiero al diseño, presentado a menudo usando diagramas, que representan la estructura de la base de datos. El modelo identifica las relaciones, atributos, claves, dominios y relaciones dentro de la base de datos. Algunos diseñadores de bases de datos crean un modelo lógico y un modelo físico. El modelo lógico se basa más en la teoría relacional y aplica los principios de normalización apropiados a los datos. El modelo físico, por otro lado, se refiere a la aplicación real, ya que los datos se almacenan en un RDBMS. Basado en el diseño lógico, el diseño físico trae la estructura de datos al mundo real de la implementación.

Pruebe esto 1-1

Normalización de datos e identificación de relaciones

Como programador principiante de SQL, es poco probable que sea responsable de la normalización de la base de datos. Sin embargo, es importante que entienda estos conceptos, al igual que es importante que entienda las clases de tratos que existen entre relaciones. La normalización y las relaciones, como las relaciones mismas, ayudan a proporcionar la base sobre la que se construye SQL. Como resultado, el ejercicio se enfoca en el proceso de normalización de datos e identificación de las relaciones entre relaciones. Para completar el ejercicio, sólo necesitará papel y lápiz en el cual esbozar el modelo de datos.

Paso a paso

1. Revise la relación en la siguiente ilustración:

ID_CD	NOMBRE_CD	CATEGORIA
99301	Famous Blue Raincoat	Folk, Pop
99302	Blue	Folk, Pop
99304	Past Light	New Age
99305	Kojiki	New Age, Classical
99306	That Christmas Feeling	Christmas, Classics
99307	Patsy Cline: 12 Greatest Hits	Country, Pop, Classics

2. Identifique cualquier elemento que no se ajuste a las tres formas normales. Encontrará que el atributo CATEGORIA contiene más de un valor por tupla, que viola la primera forma normal.
3. Normalice los datos de acuerdo a las formas normales. Projete el modelo de datos que incluya las relaciones, atributos y tuplas apropiadas. Su modelo incluirá tres tablas, una para la lista de CD, otra para la lista de las categorías de música (por ejemplo, Pop), y otra que asocie los CD con las apropiadas categorías de música. Vea el archivo en línea Try_This_01-1a.jpg para un ejemplo de cómo el modelo de datos puede verse (en inglés).
4. En la ilustración que dibujó, identifique las relaciones entre la relaciones. Recuerde que cada CD se puede asociar con una o más categorías, y cada categoría se puede asociar con cero, uno o más CD. Vea el archivo en línea Try_This_01-1b.jpg para ver las relaciones entre relaciones (en inglés).

Resumen de Pruebe esto

Los modelos de datos suelen ser más específicos que las ilustraciones mostradas en el ejercicio. Las relaciones y claves se marcan claramente con símbolos que se ajustan a un tipo particular de sistema de modelado de datos, y las relaciones muestran sólo los atributos, pero no las tuplas. Sin embargo, para efectos de este capítulo, es suficiente con que tenga los conocimientos básicos de normalización y relaciones entre relaciones. El ejercicio se entiende como un forma para que entienda mejor esos conceptos y cómo se aplican al modelo relacional.

Aprenda acerca de SQL

Ahora que tiene un conocimiento fundamental del modelo relacional, es el momento para introducirlo a SQL y sus características básicas. Como recordará de la sección “Entienda las bases de datos relacionales” vista anteriormente en este capítulo, SQL se basa en el modelo relacional, aunque no se trate de una aplicación exacta. Mientras el modelo relacional proporciona las bases teóricas de la base de datos relacional, es el lenguaje SQL el que apoya la aplicación física de esa base de datos.

SQL, el lenguaje relacional casi universalmente aplicado, es diferente de otros lenguajes computacionales como C, COBOL y Java, los cuales son de procedimiento. Un lenguaje de procedimiento define *cómo* las operaciones de una aplicación deben realizarse y el orden en el cual se realizan. Un lenguaje de no procedimiento, por otro lado, se refiere a los resultados de una operación; el entorno fundamental del software determina cómo se procesan las operaciones. Esto no quiere decir que SQL respalda a la funcionalidad de no procedimiento. Por ejemplo, los procedimientos almacenados, agregados a varios productos RDBMS hace algunos años, son parte del estándar SQL:2006 y proporciona capacidades parecidas a procedimiento. (Los procedimientos almacenados se analizan en el capítulo 13.) Muchos de los proveedores de RDBMS añadieron extensiones a SQL para proporcionar esas capacidades de procedimiento, como Transact-SQL encontrado en Sybase y Microsoft SQL Server y PL/SQL encontrado en Oracle.

SQL aún carece de muchas de las capacidades básicas de programación de la mayoría de los lenguajes computacionales. Por esta razón, a menudo SQL se considera como un *sublenguaje* de datos porque se utiliza con frecuencia en asociación con la aplicación de lenguajes de programación como C y Java, lenguajes que no fueron diseñados para la manipulación de datos almacenados en una base de datos. Como resultado, SQL se utiliza en conjunto con la aplicación del lenguaje para proporcionar un medio eficaz de acceder a los datos, razón por la cual se considera a SQL como un sublenguaje.

La evolución de SQL

A principios de la década de 1970, después que se publicó el artículo de E. F. Codd, IBM comenzó a desarrollar un lenguaje y un sistema de base de datos que podría usarse para la aplicación de ese modelo. Cuando se definió por primera vez, el lenguaje fue denominado Lenguaje de consulta estructurado (en inglés, SEQUEL, Structured English Query Language). Cuando se descubrió que SEQUEL era propiedad de una marca comercial de Hawker-Siddeley Aircraft Company en el Reino Unido, el nombre se cambió a SQL. Cuando se pasó la voz de que IBM estaba desarrollando

un sistema de base de datos relacional basado en SQL, otras compañías comenzaron a desarrollar sus propios productos basados en SQL. De hecho, Relational Software, Inc., ahora Oracle Corporation, lanzó el sistema de base de datos antes de que IBM lanzara el suyo al mercado. Conforme más proveedores lanzaron sus productos, SQL comenzó a surgir como el lenguaje estándar de base de datos relacional.

En 1986, el American National Standards Institute (ANSI) dio a conocer el primer estándar publicado para el lenguaje (SQL-86), el cual fue adoptado por la International Organization for Standardization (ISO) en 1987. El estándar se actualizó en 1989, 1992, 2003, 2006, y el trabajo continúa. Ha crecido con el tiempo (el estándar original estaba muy por debajo de 1 000 páginas, mientras que la versión de SQL:2006 tiene más de 3 700 páginas). El estándar se escribió en partes para permitir la publicación programada de revisiones y facilitar el trabajo paralelo por diferentes comités. La tabla 1-1 proporciona una descripción general de las partes y el estado actual de cada una.

Partes	Tema	Estado
1	SQL/Framework	Completado en 1999; revisado en el 2003, correcciones publicadas en el 2007
2	SQL/Foundation	Completado en 1986, revisado en 1999 y 2003, correcciones publicadas en el 2007
3	SQL/CLI	Completado en 1995, revisado en 1999 y en el 2003, correcciones publicadas en el 2005
4	SQL/PSM	Completado en 1996, revisado en 1999 y en el 2003, correcciones publicadas en el 2007
5	SQL/Bindings	Establecido como una parte separada en 1999, fusionado de nuevo en la parte 2 en el 2003; actualmente no existe la parte 5
6	SQL/Transaction	Proyecto cancelado; actualmente no existe la parte 6
7	SQL/Temporal	Retirado; actualmente no existe la parte 7
8	SQL/Objects and Extended Objects	Fusionado en la parte 2; no existe la parte 8
9	SQL/MED	Comenzado después de 1999, completado en el 2003, correcciones publicadas en el 2005
10	SQL/OLB	Completado como el estándar ANSI en 1998, versión ISO completada en 1999, revisión en el 2003, correcciones publicadas en el 2007
11	SQL/Schemata	Extraído de una parte separada en el 2003, correcciones publicadas en el 2007
12	SQL/Replication	Proyecto empezado en el 2000, pero posteriormente abandonado; actualmente no existe la parte 12
13	SQL/JRT	Completado como el estándar ANSI en 1999, revisión completada en el 2003, correcciones publicadas en el 2005
14	SQL/XML	Completado en el 2003, expandido en el 2006, correcciones publicadas en el 2007

Tabla 1-1 Partes del estándar SQL.

Los proveedores RDBMS lanzaron productos al mercado antes que hubiera un estándar, y muchas de las características de esos productos se aplicaron bastante diferentes, así que el estándar no pudo acomodarlos a todos cuando se desarrollaron. A menudo se llaman *proveedores por extensión*. Esto explica por qué no hay un estándar para una base de datos. Y cuando cada versión estándar de SQL se lanza, los proveedores RDBMS tienen que trabajar para incorporar el nuevo estándar en sus productos. Por ejemplo, los procedimientos almacenados y activadores son nuevos en el estándar SQL:1999, pero fueron implementados en RDBMS por muchos años. SQL:1999 simplemente estandarizó el lenguaje utilizado para implementar funciones que ya existían.

NOTA

Aunque se analizarán los procedimientos almacenados en el capítulo 13 y los activadores en el capítulo 14, daré una definición rápida de cada uno. Un procedimiento almacenado es un conjunto de instrucciones de SQL almacenadas como un objeto en el servidor de una base de datos al que un cliente puede recurrir simplemente llamando al procedimiento. Un activador es similar a un procedimiento almacenado en el sentido que es un conjunto de instrucciones de SQL almacenadas como un objeto en el servidor de una base de datos. Sin embargo, en lugar de que un cliente recurra al activador, éste es invocado automáticamente cuando se produce un evento predefinido, tales como la inserción o la actualización de datos.

Objeto de modelo relacional

El lenguaje SQL se basa en el modelo relacional, y hasta SQL-92, también el estándar SQL. Sin embargo, comenzando con SQL:1999, el estándar SQL se extendió más allá del modelo relacional puro para incluir construcciones orientadas a objetos en el lenguaje. Estas construcciones se basan en los conceptos inherentes de *programación orientada a objetos*, una programación metodológica que define colecciones autónomas de estructura de datos y rutinas (llamadas *objetos*). En los lenguajes orientados a objetos como Java y C++, los objetos interactúan entre sí de manera que permiten al lenguaje abordar problemas complejos que no serían fáciles de resolver en lenguajes tradicionales.

Con la llegada de la programación orientada a objetos (junto con los avances tecnológicos en el hardware y software y la creciente complejidad de aplicaciones) se hizo cada vez más evidente que un lenguaje puramente relacional era insuficiente para satisfacer las demandas del mundo real. De preocupación específica fue el hecho que SQL no podía respaldar tipos de datos complejos y definidos por el usuario ni la extensibilidad requerida para aplicaciones más complejas.

Impulsados por la competencia natural de la industria, los proveedores RDBMS se encargaron de aumentar sus productos e incorporar la funcionalidad orientada a objetos en sus sistemas. El estándar SQL:2006 sigue el ejemplo y extiende el modelo relacional con capacidades orientadas a objetos, como métodos, encapsulación, y tipos de datos complejos y definidos por el usuario, lo que hace a SQL un lenguaje de base de datos *relacional a objeto*. Como se muestra en la tabla 1-1, la parte 14 (SQL/XML) se amplió considerablemente y se reeditó con SQL:2006, y todas las demás partes se tomaron de SQL:2003.

La conformidad con SQL:2006

Una vez que SQL se estandarizó, se desprende que el estándar también define lo que se tomó para una aplicación de SQL (un producto RDBMS) considerado en conformidad al estándar. Por ejem-

pló, el estándar SQL-92 proporcionó tres niveles de conformidad: entrada, intermedio y pleno. Los RDBMS más populares alcanzaron sólo el nivel de entrada de conformidad. Debido a esto, SQL:2006 tomó un enfoque diferente al establecer las normas de conformidad. Para que un producto esté en conformidad con SQL:2006, debe apoyar el nivel de conformidad Core SQL. Core SQL en el estándar SQL:2006 se define en conformidad a la parte 2 (SQL/Foundation) y la parte 11 (SQL/Schemata) del estándar.

Además del nivel de conformidad Core SQL, los proveedores demandan la conformidad en cualquier parte cumpliendo con los requisitos mínimos de conformidad por su parte.

NOTA

Puede encontrar información sobre el estándar SQL:2006 comprando una copia del documento(s) estándar apropiado publicado por ANSI e ISO. El estándar se divide en nueve documentos (una parte por documento). El primer documento (ANSI/ISO/IEC 9075-1:2003) incluye la descripción general de las nueve partes. El sufijo del nombre de cada documento contiene el año de publicación, y las diferentes partes tienen distintos años de publicación debido a que las partes fueron actualizadas y publicadas independientemente por distintos comités. Como puede ver en la tabla 1-1, la parte 1 fue publicada en el 2003 y, de hecho, sólo la parte 14 lleva una fecha de publicación del 2006 (todas las demás partes se publicaron en el 2003). Puede adquirir estos documentos en línea en ANSI Electronic Standards Store (<http://webstore.ansi.org/>), NCITS Standards Store (<http://www.techstreet.com/ncitsgate.html>) o ISO Store (<http://www.iso.org/iso/store.htm>). En el sitio de la ANSI observe que hay dos variantes de cada documento con un contenido esencialmente idéntico, llamado INCITS/ISO/IEC 9075 e ISO/IEC 9075. Las variantes ISO/IEC cuestan entre 139 y 289 dólares por documento, mientras que las variantes INCITS/ISO/IEC cuestan entre 30 dólares por documento. ISO Store tiene el conjunto de documento disponible en un cómodo CD por 356 francos suizos (alrededor de 350 dólares). Obviamente, los precios están sujetos a cambio en cualquier momento. También están disponibles las correcciones sin cargo extra, llamadas "Correcciones técnicas". Como se muestra en la tabla 1-1, se corrigieron tres partes publicadas en 2005, y también se corrigieron las otras seis partes publicadas en 2007.

Tipos de instrucciones de SQL

Aunque SQL se considera un sublenguaje debido a su naturaleza de no procesamiento, aun así es un lenguaje completo que le permite crear y mantener objetos en una base de datos, asegurar esos objetos y manipular la información dentro de los objetos. Un método común usado para categorizar las instrucciones SQL es dividir las de acuerdo con las funciones que realizan. Basado en este método, SQL se separa en tres tipos de instrucciones:

- **Lenguaje de definición de datos (DDL, Data Definition Language)** Las instrucciones DDL se usan para crear, modificar o borrar objetos en una base de datos como tablas, vistas, esquemas, dominios, activadores, y almacenar procedimientos. Las palabras clave en SQL más frecuentemente asociadas con las instrucciones DDL son CREATE, ALTER y DROP. Por ejemplo, se usa la instrucción CREATE TABLE para crear una tabla, la instrucción ALTER TABLE para modificar las características de una tabla, y la instrucción DROP TABLE para borrar la definición de la tabla de la base de datos.

- **Lenguaje de control de datos (DCL, Data Control Language)** Las instrucciones DCL permiten controlar quién o qué (un usuario en una base de datos puede ser una persona o un programa de aplicación) tiene acceso a objetos específicos en la base de datos. Con DCL, puede otorgar o restringir el acceso usando las instrucciones GRANT o REVOKE, los dos comandos principales en DCL. Las instrucciones DCL también permiten controlar el tipo de acceso que cada usuario tiene a los objetos de una base de datos. Por ejemplo, puede determinar cuáles usuarios pueden ver un conjunto de datos específico y cuáles usuarios pueden manipular esos datos.
- **Lenguaje de manipulación de datos (DML, Data Manipulation Language)** Las instrucciones DML se usan para recuperar, agregar, modificar o borrar datos almacenados en los objetos de una base de datos. Las palabras clave asociadas con las instrucciones DML son SELECT, INSERT, UPDATE y DELETE, las cuales representan los tipos de instrucciones que probablemente son más usadas. Por ejemplo, puede usar la instrucción SELECT para recuperar datos de una tabla y la instrucción INSERT para agregar datos a una tabla.

La mayoría de las instrucciones SQL que se utilizan caen perfectamente en una de estas categorías, y estas instrucciones se analizarán durante el resto del libro.

NOTA

Hay varias formas de clasificar las instrucciones además de la manera en que se clasifican en la lista anterior. Por ejemplo, se pueden clasificar de acuerdo a cómo se ejecutan o si pueden o no ser incrustadas en un lenguaje de programación estándar. El estándar SQL proporciona diez categorías amplias basadas en funciones. Sin embargo, se usa el método anterior ya que se utiliza comúnmente en la documentación relacionada con SQL, y porque es una manera simple de proporcionar una buena visión general de la funciones inherentes a SQL.

Tipos de ejecución

Además de definir cómo se usa el lenguaje, el estándar SQL proporciona detalles de cómo las instrucciones SQL son ejecutadas. Este método de ejecución, conocido como *estilos de unión*, no sólo afecta la naturaleza de la ejecución, sino también determina cuáles instrucciones, como mínimo, deben ser soportadas por un estilo de unión particular. El estándar define cuatro métodos de ejecución:

- **Invocación directa** Mediante el uso de este método, puede comunicarse directamente desde una aplicación de usuario, como iSQL*Plus en Oracle o Management Studio en Microsoft SQL Server, en la base de datos. (La aplicación de usuario y la base de datos pueden estar en la misma computadora, pero a menudo no lo están.) Simplemente introduzca su consulta en la ventana de la aplicación y ejecute la instrucción SQL. Los resultados de su consulta se le devolverán tan rápido como el poder del procesador y las limitaciones de la base de datos lo permitan. Ésta es una forma rápida de comprobar datos, verificar conexiones y ver los objetos en una base de datos. Sin embargo, las directrices del estándar SQL sobre la invocación directa son bastante mínimas; por lo tanto, los métodos utilizados y los estándares SQL respaldados pueden variar ampliamente de un producto a otro.

- **SQL incrustado** En este método, las instrucciones SQL están codificadas (incrustadas) directamente en el lenguaje de programación anfitrión. Por ejemplo, las instrucciones SQL se pueden incrustar en el código C de la aplicación. Antes que el código se compile, un pre-procesador analiza las instrucciones SQL y las desglosa desde el código C. El código SQL se convierte en una forma que RDBMS puede entender, y el código C restante se compila como lo haría normalmente.
- **Unión de módulo** Este método permite crear bloques de instrucciones SQL (módulos) que están separados del lenguaje de programación anfitrión. Una vez que el módulo es creado, es una combinación entre una aplicación y un vinculador. Un módulo contiene, entre otras cosas, procedimientos, y son los procedimientos los que contienen las instrucciones SQL reales.
- **Interfaz convocatoria a nivel (CLI, Call-level interface)** Una CLI permite invocar instrucciones SQL a través de una interfaz mediante la aprobación de instrucciones SQL como valores argumentativos para las subrutinas. Las instrucciones no están precompiladas como en el SQL incrustado y la Unión de módulo. En lugar de eso, son ejecutadas directamente por los RDBMS.

La invocación directa, aunque no es el método utilizado más comúnmente, es el que será usado principalmente para los ejemplos y ejercicios en este libro, ya que apoya la presentación de las consultas *ad hoc* en la base de datos y genera resultados inmediatos. Sin embargo, SQL incrustado es actualmente el método utilizado más comúnmente en las aplicaciones de negocios. Se analizará este método, así como el unión de módulo y CLI, con mayor detalle en el capítulo 17.

Pregunta al experto

- P:** Se afirma que, para que un RDBMS esté en conformidad con el estándar SQL:2006, debe cumplir con Core SQL. ¿Hay algún requisito adicional que un producto debe cumplir?
- R:** Sí. Además de Core SQL, un RDBMS debe respaldar a ambos el SQL incrustado o la unión de módulo. La mayoría de los productos respaldan sólo al SQL incrustado, algunos respaldan a ambos. El estándar SQL no requiere productos RDBMS para respaldar la invocación directa o CLI, aunque las mayoría lo hace.
- P:** ¿Cuáles son las diez categorías usadas por el estándar SQL:2006 para clasificar las instrucciones SQL?
- R:** El estándar SQL clasifica las instrucciones en las siguientes categorías: esquema, datos, cambio de datos, operación, conexión, control, sesión, diagnósticos, dinámico e instrucción incrustada de excepción. Tenga en cuenta que estas clasificaciones no son más que una herramienta que puede utilizar para comprender mejor el alcance del lenguaje y los conceptos básicos. En última instancia, son las instrucciones SQL (y lo que pueden hacer) lo que es importante.

Utilice un sistema de gestión de base de datos relacional

A lo largo de este capítulo, al examinar el modelo relacional y SQL, se mencionaron a menudo los RDBMS y la forma en que utilizan el estándar SQL como base de sus productos. Un *sistema de gestión de base de datos relacional* es un programa o conjunto de programas que almacenan, administran, recuperan, modifican y manipulan información en una o más bases de datos relacionales. Oracle, Microsoft SQL Server, IBM's DB2 y el producto de prueba MySQL son todos ejemplos de RDBMS. Estos productos, como otros RDBMS, permiten interactuar con la información almacenada en el sistema. Aunque no es necesario que un RDBMS se base en SQL, la mayoría de los productos en el mercado están basados en SQL y tratan de ajustarse al estándar SQL. Como mínimo, estos productos demandan el nivel de entrada de conformidad con el estándar SQL-92, y ahora se trabaja para Core SQL en conformidad con SQL:2006.

Además del cumplimiento con los estándares SQL, la mayoría de los RDBMS apoyan otras características, como instrucciones SQL adicionales, productos basados en herramientas administrativas, y aplicaciones de interfaz gráfica de usuario (GUI) que permiten la consulta y manipulación de información, la gestión de objetos en una base de datos, y la administración del sistema y su estructura. Los tipos de funcionalidad aplicados y los métodos utilizados para proporcionar la funcionalidad pueden variar ampliamente de un producto a otro. Conforme las bases de datos crecen, se vuelven más complicadas y se distribuyen en más áreas, los productos RDBMS utilizados para la gestión de las bases de datos se hacen más complejos y robustos, satisfaciendo las demandas del mercado, así como aplicación de nuevas y más sofisticadas tecnologías.

El estándar SQL frente a las implementaciones de producto

El centro de cualquier RDBMS basado en SQL es, por supuesto, el propio SQL. Sin embargo, el lenguaje utilizado no es SQL puro. Cada producto extiende su lenguaje con el fin de implementar las características definidas por el proveedor y mejorar la funcionalidad basada en SQL. Además, una serie de productos RDBMS lo fabricaron para el mercado antes de que hubiera un estándar. En consecuencia, cada proveedor respalda una variación ligeramente diferente de SQL, lo que significa que el lenguaje utilizado en cada producto tiene una aplicación específica. Por ejemplo, SQL Server utiliza Transact-SQL, que incluye tanto a SQL como a las extensiones del proveedor para proporcionar las instrucciones de procedimiento necesarias para los activadores y los procedimientos almacenados. Por otro lado, Oracle proporciona instrucciones de procedimiento en un componente de producto separado llamado PL/SQL. Como resultado, las instrucciones SQL que se proporcionan en el libro pueden ser ligeramente diferentes en la aplicación del producto que utiliza.

A lo largo del libro se utilizará SQL puro en la mayoría de los ejemplos y ejercicios. Sin embargo, me doy cuenta de que, como un programador principiante de SQL, su interés principal es la aplicación de SQL en el mundo real. Por esa razón, a veces usaré SQL Server (con Transact-SQL) u Oracle (con PL/SQL) para demostrar o aclarar un concepto particular que no puede explicarse totalmente por medio de SQL puro únicamente.

Una de las ventajas de utilizar un producto como Oracle o SQL Server es que ambos respaldan la invocación directa a través de una aplicación GUI de usuario. SQL Server usa la interfaz Management Studio, que se muestra en la figura 1-7. La interfaz GUI hace posible crear una consulta específica en SQL, presentarla a DBMS para la transformación y ver los resultados, lo que le permite aplicar lo que se aprende en el libro en un entorno real de SQL. Oracle tiene varias soluciones para un GUI de usuario, incluyendo el interfaz basado en la web *iSQL** Plus, mostrado en la figura 1-8.

Además de las interfaces GUI, la mayoría de los productos incluyen una interfaz de línea de comandos que se utiliza en las terminales más antiguas que no tienen capacidad gráfica. Estas interfaces también son útiles para la ejecución de instrucciones SQL y por las conexiones de acceso telefónico donde las interfaces gráficas son demasiado lentas. En la figura 1-9 se muestra la interfaz de línea de comandos por el cual MySQL se ejecuta en Microsoft Windows.

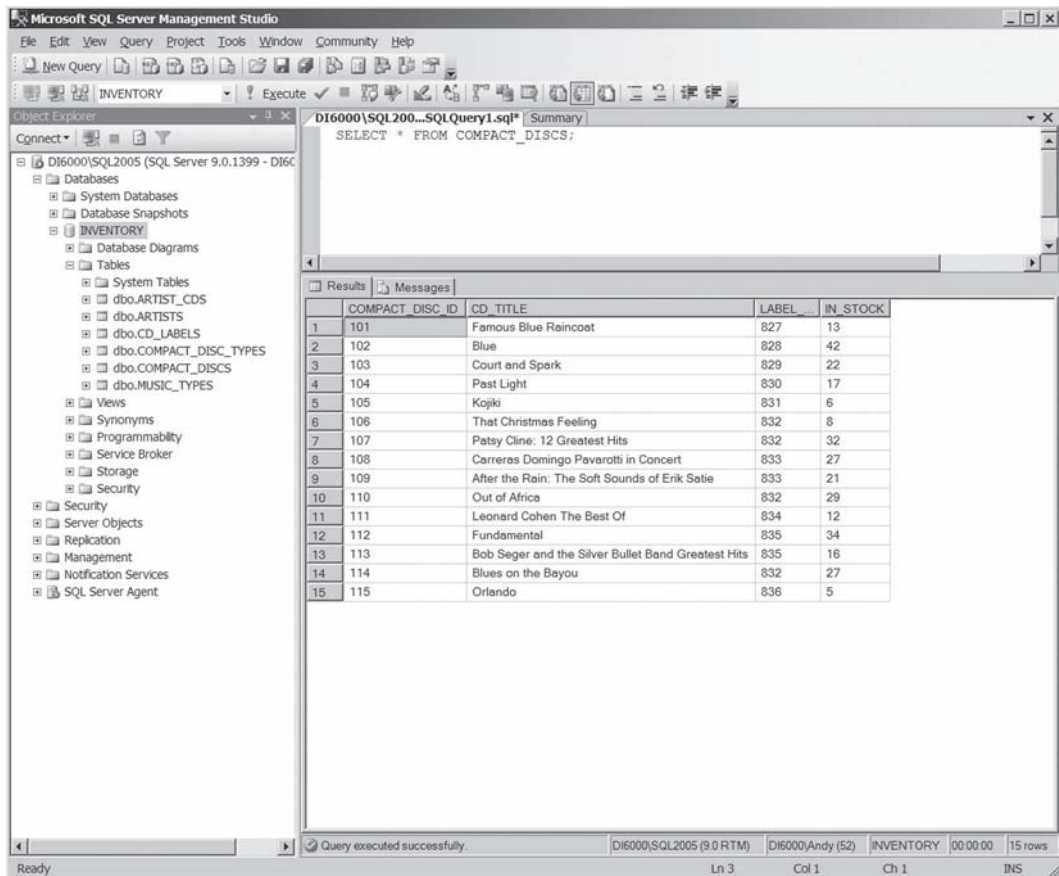


Figura 1-7 Uso de SQL Server Management Studio.

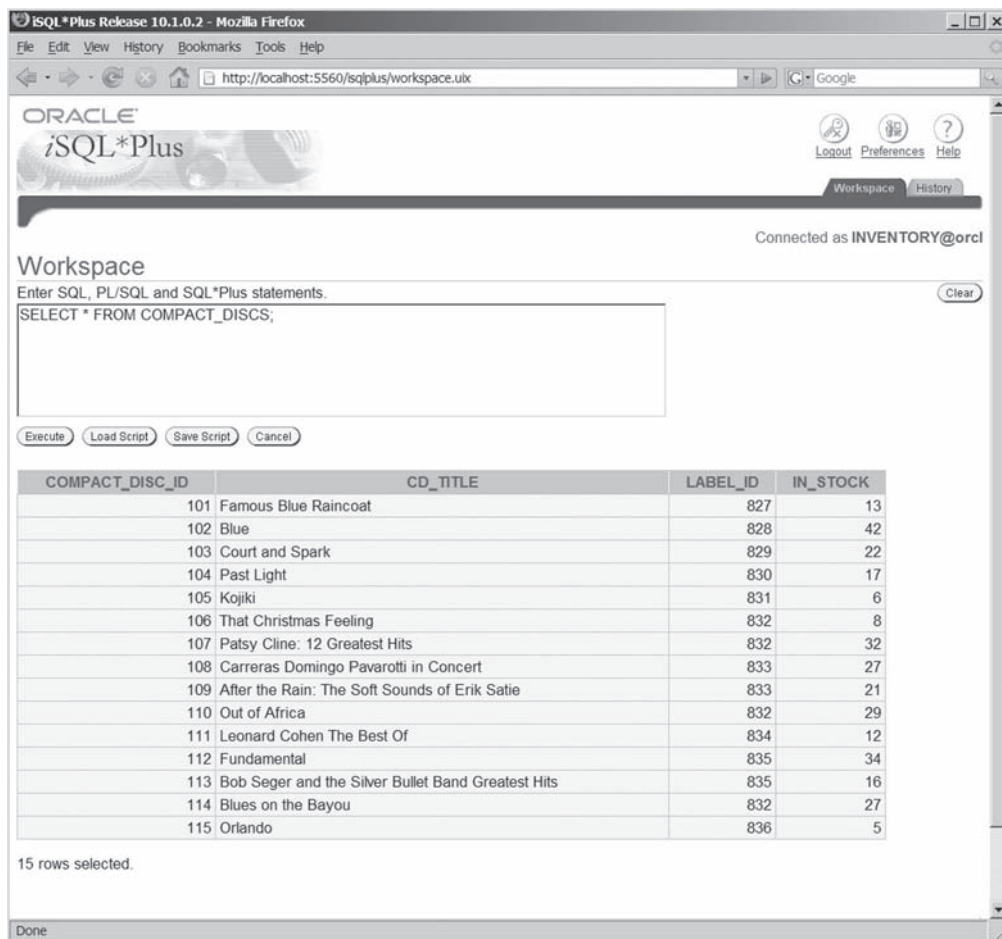
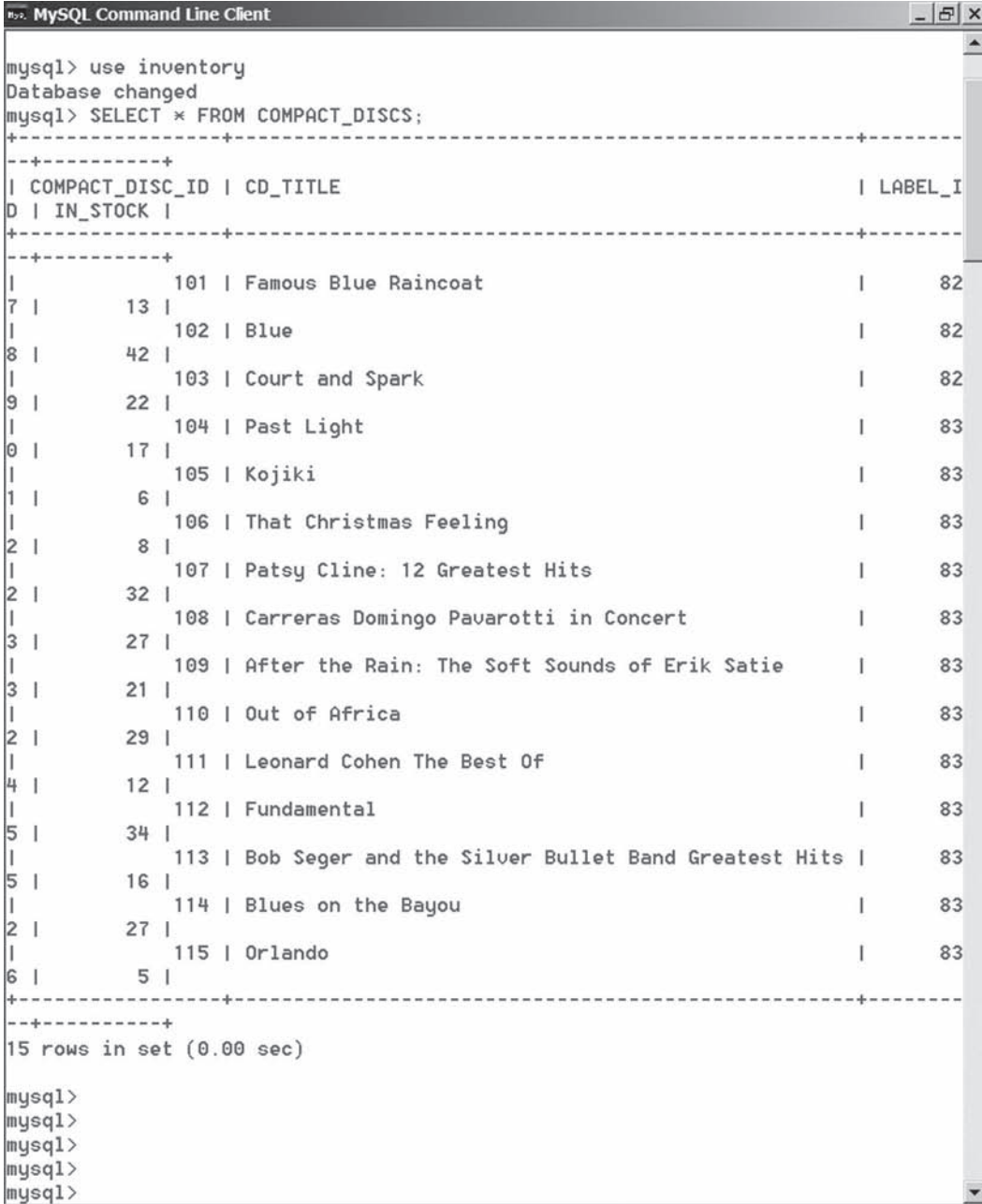


Figura 1-8 Uso de Oracle iSQL*Plus.

El uso de estos productos no implica que se apoye a alguno de ellos por encima de cualquier otro producto comercial (como Sybase o Informix) o los productos de prueba (como PostgreSQL), y de hecho, le animamos a utilizar cualquier RDBMS que tenga disponible, suponiendo que da soporte a gran parte de la funcionalidad que se analiza en este libro. Sin embargo, se eligen SQL Server y Oracle para la mayoría de los ejemplos en el libro, ya que se demostrará cómo se aplica SQL en el mundo real y cómo SQL puede diferir de una aplicación específica de una versión del lenguaje, y esos dos productos suministran el medio para llevarlo a cabo. Tenga en cuenta que, para que pueda tener una plena comprensión de SQL y pueda utilizarlo en varios productos RDBMS, necesitará entender tanto el estándar SQL como el lenguaje que se aplica a los productos que van a utilizar.



```

mysql> use inventory
Database changed
mysql> SELECT * FROM COMPACT_DISCS;
+-----+-----+-----+
| COMPACT_DISC_ID | CD_TITLE                                     | LABEL_I |
| IN_STOCK |
+-----+-----+-----+
|          101 | Famous Blue Raincoat                       |         82 |
7 |          13 |
|          102 | Blue                                       |         82 |
8 |          42 |
|          103 | Court and Spark                           |         82 |
9 |          22 |
|          104 | Past Light                               |         83 |
0 |          17 |
|          105 | Kojiki                                    |         83 |
1 |           6 |
|          106 | That Christmas Feeling                    |         83 |
2 |           8 |
|          107 | Patsy Cline: 12 Greatest Hits              |         83 |
2 |          32 |
|          108 | Carreras Domingo Pavarotti in Concert     |         83 |
3 |          27 |
|          109 | After the Rain: The Soft Sounds of Erik Satie |         83 |
3 |          21 |
|          110 | Out of Africa                             |         83 |
2 |          29 |
|          111 | Leonard Cohen The Best Of                 |         83 |
4 |          12 |
|          112 | Fundamental                               |         83 |
5 |          34 |
|          113 | Bob Seger and the Silver Bullet Band Greatest Hits |         83 |
5 |          16 |
|          114 | Blues on the Bayou                        |         83 |
2 |          27 |
|          115 | Orlando                                   |         83 |
6 |           5 |
+-----+-----+-----+
15 rows in set (0.00 sec)

mysql>
mysql>
mysql>
mysql>
mysql>

```

Figura 1-9 Uso de la interfaz de línea de comandos MySQL.

Pruebe esto 1-2 Conexión a una base de datos

Aunque este libro se centra principalmente en SQL puro, para probar los ejemplos y hacer la mayoría de los ejercicios, necesitará acceder a un RDBMS para ejecutar las instrucciones SQL. Como resultado, una de las primeras cosas que debe hacer es asegurarse que pueda acceder a un entorno SQL. Este ejercicio le ayudará a hacerlo; sin embargo, a diferencia de la mayoría de los otros ejercicios en el libro, éste requerirá un mayor esfuerzo de su parte para buscar recursos fuera de la obra para establecerse con un RDBMS que le permita invocar las instrucciones SQL directamente. Con este fin, el ejercicio intenta que empiece, pero debe utilizar su propia iniciativa para garantizar que tiene un entorno en el que se sienta cómodo trabajando.

Paso a paso

1. Identifique el RDBMS que va a utilizar para los ejercicios del libro. Quizás haya un sistema con el cual ya está familiarizado o uno que esté disponible en su entorno de trabajo. Si no tiene ninguno disponible en su trabajo y no está listo para comprar un producto, revise en línea para ver cuál está disponible. La mayoría de los proveedores RDBMS ofrecen una versión gratuita básica de uso restringido de sus productos, a menudo denominada “Express Edition” (la edición sencilla); otros ofrecen una versión de prueba que puede usarse gratis por un periodo limitado. Por ejemplo, puede descargar Oracle Express Edition para Linux o Windows en <http://www.microsoft.com/sql/editions/express/>. (Necesitará una conexión a Internet de alta velocidad para descargar archivos de gran tamaño.)

Si prefiere un producto de prueba, MySQL es una opción popular. De hecho, tanto Yahoo como Google utilizan MySQL ampliamente en una configuración conocida como LAMP (Linux, Apache, MySQL y PHP). Puede descargar MySQL Community Server gratis de <http://dev.mysql.com/downloads/mysql>. Por otro lado, MySQL Enterprise Server es una edición con una tarifa respaldada por el proveedor.

Antes que se decida por un producto en particular, haga las investigaciones necesarias para asegurarse de que admite la invocación directa, preferiblemente a través de una aplicación GUI, y pueda ejecutarse en su computadora. También compruebe qué tanto del estándar SQL:2006 respalda y revise los acuerdos de licencia para asegurarse de que está en el cumplimiento. Si un sistema está disponible en su trabajo, asegúrese de hablar con los administradores de base de datos y de red para determinar qué servidor puede usar, si puede y cómo debe descargar una copia, y cómo hacer la conexión con el servidor SQL. A menudo se necesita una cuenta para conectarse a la RDBMS, por lo que si éste es el caso, averigüe qué nombre de usuario y contraseña debe usar.

2. Una vez que establezca cuál RDBMS usará, instálelo en su equipo. Si se conecta a un sistema a través de la red, tendrá que instalar sólo las herramientas de clientes en su equipo local.
3. Abra el cliente GUI que le permita invocar directamente las instrucciones SQL. Cuando abra el GUI, es posible que se le pida un nombre de usuario y contraseña. Cuándo y si se le pide puede variar dependiendo del producto que utiliza, si se conecta a través de la red, si el RDBMS se configura como un sistema autónomo, y otras variables específicas del producto. Además, un producto como SQL Server ofrece seguridad integrada con el sistema operativo, por lo que es posible que sólo se le pida un nombre de servidor.

(continúa)

4. Ejecute la instrucción `SELECT` en la aplicación de entrada de la ventana. Me doy cuenta que aún no se cubren las instrucciones `SELECT`, pero la sintaxis básica es relativamente fácil:

```
SELECT * FROM <tabla>
```

El marcador de posición `<tabla>` debe sustituirse con el nombre de una tabla en una base de datos existente.

El propósito de este ejercicio es simplemente comprobar que tiene conectividad con los datos almacenados en el RDBMS. La mayoría de los productos incluyen información de muestra, y esa información es con la que intentará conectarse. Compruebe la documentación del producto o consulte con el administrador de bases de datos para verificar si existe una base de datos que pueda acceder.

Si trabaja en Oracle y los esquemas de muestra están instalados, puede ejecutar las siguientes instrucciones:

```
SELECT * FROM scott.emp;
```

Para ejecutar las instrucciones, escríbalas en la ventana de entrada de `iSQL*Plus` y luego dé un clic en `EXECUTE`.

Si trabaja en SQL Server con la información muestra instalada, puede ejecutar la siguiente instrucción:

```
USE pubs
SELECT * FROM employee
```

Para ejecutar la instrucción, escríbala en la ventana de entrada de SQL Server Management Studio y luego dé un clic en `EXECUTE`.

Si trabaja en el interfaz de línea de comandos MySQL, simplemente escriba la instrucción SQL y oprima `ENTER`.

Una vez que ejecute la instrucción, los resultados de la consulta aparecerán en la ventana de salida. En este momento, no se preocupe por el significado de cada palabra en la instrucción SQL o con los resultados de la consulta. Su única preocupación es asegurarse que todo funcione. Si no puede ejecutar la instrucción, consulte con el administrador de bases de datos o la documentación del producto.

5. Cierre la aplicación GUI sin guardar la consulta.

Resumen de Pruebe esto

Como se dijo al principio del ejercicio, es diferente de la mayoría de los demás en el libro debido a que está básicamente por su propia cuenta para establecer la conectividad con el RDBMS. De nuevo, esto se debe a que SQL es un lenguaje estándar, independiente de las implementaciones RDBMS, y las cuestiones específicas de los proveedores están, en su mayor parte, más allá del alcance de este libro. Además, los métodos utilizados para conectarse a una base de datos, las herramientas disponibles para hacer esas conexiones y la forma en la que un RDBMS está configurado varía de un producto a otro, de un entorno a otro e incluso de un sistema operativo a otro. Sin embargo, el tiempo que tome ahora para investigar qué producto utilizar y para asegurarse de que puede conectar la información existente en una base de datos será de valor incalculable conforme aplique información analizada en el resto del libro.

✓ Autoexamen Capítulo 1

1. ¿Qué es una base de datos?
2. ¿Cuál de los siguientes objetos conforma una relación?
 - A Tipos de datos
 - B Tuplas
 - C Atributos
 - D Formas
3. Un(a) _____ es un conjunto de datos cuyos valores conforman una instancia de cada uno de los atributos definidos para esa relación.
4. ¿Cuáles son las diferencias entre la primera forma normal y la segunda forma normal?
5. Una relación está en la tercera forma normal si está en la segunda forma normal y si cumple con las demás directrices de esa forma. ¿Cuáles son esas directrices?
6. ¿Cuáles son los tres principales tipos de relaciones soportados por una base de datos relacional?
7. En el modelo de datos, hay dos relaciones asociadas cada una entre sí por una relación varias a varias. ¿Cómo se implementa físicamente esta relación en una base de datos relacional?
8. ¿Cómo se diferencia SQL de los lenguajes de programación como C, COBOL y Java?
9. ¿Qué factores contribuyeron a que el estándar SQL:2006 incorporara capacidades orientadas a objetos?
10. ¿Qué nivel de conformidad debe soportar un RDBMS para poder cumplir con SQL:2006?
 - A Entrada
 - B Core
 - C Pleno
 - D Intermedio
11. ¿Cuáles son las diferencias entre las instrucciones DDL y DML?
12. ¿Qué método de ejecución de instrucciones SQL usaría si deseara comunicarse directamente con una base de datos SQL desde una aplicación de usuario?
13. ¿Cuáles son los cuatro métodos que soporta el estándar SQL:2006 para la ejecución de las instrucciones SQL?
14. ¿Qué es un sistema de gestión de base de datos relacional?
15. ¿Cuál es un ejemplo de un RDBMS?

Capítulo 2

Trabajo con el entorno SQL

Habilidades y conceptos clave

- Entienda el entorno SQL
 - Entienda los catálogos SQL
 - Nombrado de objetos en un entorno SQL
 - Creación de un esquema
 - Creación de una base de datos
-

En el capítulo 1 se analizó la teoría relacional, SQL, y los sistemas de gestión de base de datos relacional (RDBMS). En este capítulo deseo llevar este análisis un paso más allá e introducir el entorno SQL, como se define en el estándar SQL:2006. El entorno SQL proporciona la estructura en la que se ejecuta SQL. Dentro de esta estructura puede utilizar las instrucciones SQL para definir los objetos en una base de datos y almacenar información en dichos objetos. Sin embargo, antes de comenzar a escribir instrucciones SQL debe tener un entendimiento básico de los fundamentos sobre los cuales el entorno SQL se construye para que pueda aplicar esta información en todo el resto del libro. De hecho, puede resultar útil hacer referencia a este capítulo a menudo como ayuda para adquirir una comprensión conceptual del entorno SQL y cómo se relaciona con los elementos de SQL acerca de los cuales aprenderá en los siguientes capítulos.

Entienda el entorno SQL

El *entorno SQL* es, simplemente, la suma de todas las partes que conforman ese entorno. Cada parte, o componente, trabaja en conjunto con otros componentes para respaldar las operaciones de SQL tales como la creación y modificación de objetos, almacenamiento y consulta de información, o modificación y eliminación de datos. En conjunto, estos componentes forman un modelo en el que un RDBMS puede basarse. Esto no significa, sin embargo, que los proveedores de RDBMS se adhieren estrictamente a este modelo; cuáles componentes implementan, y cómo lo hacen se deja, en su mayor parte, a la discreción de esos proveedores. Aun así, deseo darle una visión general de la forma en que el entorno SQL se define, en términos de los distintos componentes, tal como se describen en el estándar SQL:2006.

El entorno SQL se compone de seis tipos de componentes, como se muestra en la figura 2-1. Los clientes SQL y los servidores SQL son parte de la aplicación de SQL y son, por lo tanto, subtipos de ese componente.

Observe que sólo hay un agente SQL y una aplicación de SQL, pero hay varios componentes de otros tipos, tales como catálogos y sitios. De acuerdo con SQL:2006, debe haber exactamente un agente SQL y una aplicación de SQL y cero o más módulos de cliente SQL, identificador de autorización y catálogos. El estándar no especifica cuántos sitios respalda, pero implica varios sitios.

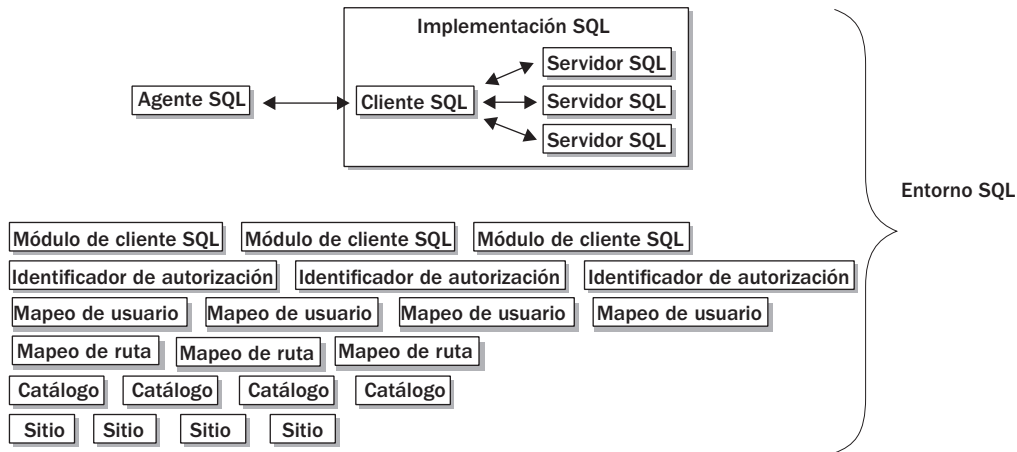


Figura 2-1 Componentes del entorno SQL.

Cada tipo de componente realiza una función específica dentro del entorno SQL. La tabla 2-1 describe los ocho tipos.

En su mayor parte, es necesario que tenga sólo una comprensión básica de los componentes que forman el entorno SQL (en términos de principios de programación de SQL). Sin embargo, uno de estos componentes (el catálogo) desempeña un papel más crítico que los otros; respecto a esto es lo que aprenderá en este libro. Por lo tanto, se cubrirá este tema con más detalle y se explicará cómo se relaciona con la gestión de datos y los objetos que sostienen dichos datos.

Tipo de componente	Descripción
Agente SQL	Cualquier estructura que haga que las instrucciones SQL se ejecuten. El agente SQL está ligado al cliente SQL dentro de la implementación de SQL.
Aplicación de SQL	Un procesador que ejecuta las instrucciones SQL en función de las necesidades del agente SQL. La aplicación de SQL incluye un cliente SQL y uno o más servidores SQL. El cliente SQL establece conexiones SQL con los servidores SQL y mantiene los datos relativos a las interacciones con el agente SQL y los servidores SQL. Un servidor SQL gestiona la sesión SQL que tiene lugar durante la conexión SQL y ejecutará instrucciones SQL recibidas del cliente SQL.

Tabla 2-1 Tipos de componentes respaldados en un entorno SQL.

Tipo de componente	Descripción
Módulo de cliente SQL	Una colección de instrucciones SQL que se escriben por separado de su lenguaje de programación de aplicación pero que pueden ser llamados dentro de ese lenguaje. Un módulo de cliente SQL contiene cero o más procedimientos invocados externamente; cada procedimiento consiste en una única instrucción SQL. Un módulo de cliente SQL reside dentro del entorno SQL y se procesa por la aplicación SQL, a diferencia de SQL incrustado, el cual se escribe dentro de la aplicación de lenguaje de programación y precompila antes de que el lenguaje de programación se elabore. Los módulos de cliente SQL se analizarán con mayor detalle en el capítulo 17.
Identificador de autorización	Un identificador representa a un usuario o un rol al que se le concede privilegios de acceso a objetos o información dentro del entorno SQL. Un usuario es una cuenta individual de seguridad que puede representar a un individuo, una aplicación o un servicio del sistema. Un <i>rol</i> es un conjunto de privilegios predefinidos que se asignan a un usuario o a otro rol. Se analizarán los identificadores autorizados, usuarios y roles en el capítulo 6.
Mapeo de usuario	Un mapeo de usuario equipara a un identificador de autorización con un descriptor de servidor exterior.
Mapeo de rutas	Un mapeo de rutas equipara a una ruta invocada en SQL con un descriptor de servidor exterior.
Catálogo	Un grupo de esquemas reunidos en un nombre definido. Cada catálogo contiene un esquema de información, que incluye los descriptores de una serie de objetos de esquema. El propio catálogo proporciona una estructura jerárquica para organizar los datos dentro de los esquemas. (Un <i>esquema</i> es básicamente un contenedor de objetos como tablas, vistas y dominios, los cuales se analizarán con mayor detalle en la siguiente sección, “Entienda los catálogos SQL.”)
Sitio	Un grupo de tablas base que contienen datos SQL, tal como lo describe el contenido de los esquemas. Se puede pensar en estos datos como “la base de datos”, pero tenga en cuenta que el estándar SQL no incluye una definición del término “base de datos”, ya que tiene muchos significados diferentes.

Tabla 2-1 Tipos de componentes respaldados en un entorno SQL (*continuación*).

Entienda los catálogos SQL

En la sección anterior, “Entienda el entorno SQL”, se expuso que un entorno SQL es la suma de todas las partes que lo componen. Se puede utilizar esa misma lógica para describir un *catálogo*, en el sentido de que un catálogo es una colección de esquemas, y esos esquemas, en conjunto, definen un nombre dentro del entorno SQL.

NOTA

Un espacio de nombre es una estructura de nombres que identifica los componentes relacionados en un determinado entorno. Un nombre es a menudo representado por una configuración de un árbol invertido para representar la relación jerárquica de objetos. Por ejemplo, suponga que su nombre incluye dos objetos: OBJETO_1 y OBJETO_2. Si el nombre es NOMBRE_1, los nombres completos de los objetos son NOMBRE_1.OBJETO_1 y NOMBRE_1.OBJETO_2 (o algunos como nomenclatura de configuración), lo que indica que comparten el mismo nombre.

Otra manera de ver a un catálogo es como una estructura jerárquica con el catálogo como el objeto primario y los esquemas como los objetos secundarios, como se muestra en la figura 2-2. En la parte superior de la jerarquía está el entorno SQL, que puede contener cero o más catálogos (aunque un entorno con cero catálogos no es bueno, ya que en el catálogo es donde se encuentran las definiciones de la información y los datos SQL). Los esquemas se localizan en el tercer nivel, debajo del catálogo, y los objetos de esquema están en el cuarto nivel.

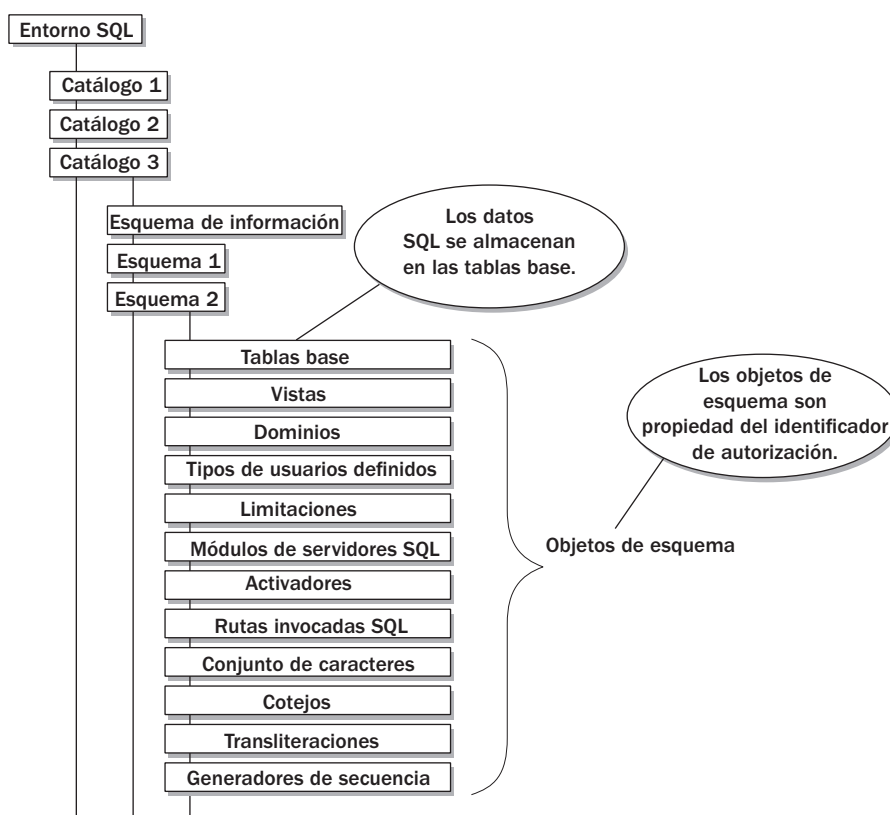


Figura 2-2 Los componentes de un catálogo.

Puede comparar las relaciones entre los objetos en un catálogo a las relaciones entre los archivos y directorios en el sistema operativo de su equipo. El catálogo se representa por un directorio raíz; los esquemas, por subdirectorios, y los objetos de esquema, por archivos dentro de subdirectorios.

Al igual que la estructura jerárquica de un sistema de archivo, la estructura de un catálogo es *lógica* por naturaleza; esto es, un sistema de archivos se presenta en una forma jerárquica (al igual que la de Windows Explorer), pero eso no significa que los archivos son almacenados jerárquicamente en el disco duro. En el mismo sentido, el catálogo de jerarquía no es más que una representación de la relación entre objetos en el entorno SQL, por lo que no implica ninguna contención física o una organización. Cómo se implementan estos objetos realmente con relación a la estructura del catálogo, y cuáles se implementan, se deja a la discreción de los proveedores de RDBMS. De hecho, el estándar SQL:2006 no define un lenguaje para la creación o eliminación de catálogos; esto también se deja a los proveedores, e incluso pocos sistemas respaldan a los catálogos.

Esquemas

Cada catálogo contiene uno o más esquemas. Un *esquema* es un conjunto de objetos relacionados que se reúnen bajo un nombre común. El esquema actúa como un contenedor de esos objetos, los que a su vez almacenan los datos SQL o realizan otras funciones con datos relacionados. Cada esquema, los objetos contenidos en el esquema y los datos SQL dentro de esos objetos son propiedad del identificador de autorización asociado con ese esquema.

A diferencia de los catálogos, los esquemas son ampliamente implementados en los productos RDBMS. Sin embargo, como con los catálogos, SQL deja la mayor parte de los detalles de implementación a los proveedores, aun cuando el estándar proporciona un lenguaje para la creación y eliminación de esquemas. Para crear un esquema se usa la instrucción `CREATE SCHEMA`, y para eliminar un esquema se usa la instrucción `DROP SCHEMA`. La creación y eliminación de esquemas se analiza con mayor detalle en la sección “Creación de un esquema”.

El tratamiento de los esquemas en un RDBMS puede variar ampliamente del estándar, y por lo tanto es importante que lea cuidadosamente la documentación del producto si desea crear un esquema en el entorno SQL. Por ejemplo, la arquitectura de la base de datos en Oracle combina el concepto de un esquema y la propiedad del identificador de autorización (cuando se crea a un usuario en Oracle, también está implícita la creación de un esquema para ese usuario). Mientras que Oracle 11g contiene la instrucción `CREATE SCHEMA` de compatibilidad con el estándar SQL, simplemente permite ejecutar un lote específico de instrucciones SQL para crear tablas y vistas y la concesión de privilegios dentro de un esquema que ya existe (es decir, uno que ya fue implícitamente creado usando la instrucción `CREATE USER`).

Esquema de información

Cada catálogo contiene un esquema especial llamado `INFORMATION_SCHEMA`. Este esquema contiene las definiciones de una serie de objetos de esquema, la mayoría de vista. Una *vista* es una tabla virtual que permite observar los datos reunidos de tablas reales. Mediante el uso de esas vistas, puede mostrar las definiciones de los objetos en ese catálogo como si se tratara de datos SQL. No puede cambiar ninguno de los datos (si lo hiciera cambiaría las definiciones objeto), pero puede mostrar la información simplemente consultando la vista correspondiente.

Como en la mayoría de características de SQL, la aplicación del esquema de información y qué funciones respalda varía de un producto a otro, aunque estas implementaciones son generalmente bastante sencillas. Por ejemplo, SQL Server 2008 incluye una vista del esquema de información llamado `INFORMATION_SCHEMA.COLUMNS`. Si consulta esta vista, los resultados incluirán una lista que contiene información acerca de cada columna accesible para el usuario actual dentro de la actual base de datos. Los resultados incluyen tanto la información como el nombre de la columna, el tipo de datos asignados a esa columna y el propietario (identificador de autorización) a quien pertenece esa columna.

Objetos de esquema

En la parte inferior del nivel del catálogo jerárquico se ubican los objetos de esquema. Los *objetos de esquema* son un conjunto de componentes relacionados que están contenidos dentro de un esquema. Éste es el nivel en el que se almacenan los datos SQL y, en consecuencia, el nivel que más preocupa a los programadores de SQL. Mediante el uso de SQL, será capaz de definir los objetos de SQL, modificar esas definiciones, y almacenar y manipular los datos de SQL dentro de los objetos. De hecho, mucho de lo que se hará en este libro de aquí en adelante tiene un impacto directo o está conectado directamente con los objetos de esquema.

El estándar SQL:2006 define 12 tipos de objetos de esquema. Estos objetos, descritos en la tabla 2-2, proporcionan las bases para el entorno SQL y la estructura de la forma en que se almacenan los datos dentro de ese entorno. Se discutirá sobre estos objetos con mayor detalle más adelante en el libro; como resultado de ello, se incluyen referencias, en su caso, a las disposiciones aplicables a los capítulos.

Esquema de objeto	Descripción
Tabla base	La unidad básica de gestión de datos en el entorno SQL. Una tabla se compone de columnas y filas y es análoga a una relación (con sus atributos y tuplas) en una teoría relacional. Cada columna se asocia con un tipo de datos y mantiene los valores que están de algún modo relacionados entre sí. Por ejemplo, una tabla de clientes contiene columnas que tienen información acerca de esos clientes, como sus nombres y direcciones. (Ver el capítulo 3.)
Vista	Una tabla virtual que se crea cuando se invoca la vista (al llamar su nombre). La tabla no existe realmente (sólo la instrucción SQL que define la tabla se almacena en la base de datos). Cuando se invoca esa instrucción, la vista toma los datos de las tablas base y muestra los resultados como si los viera de una tabla base de consulta. (Ver el capítulo 5.)
Dominio	Un objeto definido por un usuario que puede especificarse en lugar de un tipo de datos en el momento de especificar una columna (una parte del proceso de crear o alterar una definición de tabla). (Ver el capítulo 4.)

Tabla 2-2 Tipos de objetos que se definen en cada esquema.

Esquema de objeto	Descripción
Tipo definido por el usuario (UDT)	Un objeto definido por el usuario se puede especificar en lugar de un tipo de datos cuando se define una columna. SQL respalda dos tipos de UDT: distinto y estructurado. Los tipos <i>distintos</i> se basan en los tipos de datos en SQL y sus valores definidos. Los tipos <i>estructurados</i> se componen de valores de atributos, cada uno de los cuales se basan en el tipo de datos de SQL. (Ver el capítulo 3.)
Restricción	Una restricción definida en una tabla, columna o dominio que limita el tipo de datos que se inserta en el objeto de aplicación. Por ejemplo, se puede crear una restricción en una columna que restrinja los valores que se insertan en esa columna a un rango específico o una lista de números. (Ver el capítulo 4.)
Módulo de servidor SQL	Un módulo que contiene rutinas invocadas en SQL. Un <i>módulo</i> es un objeto que contiene instrucciones SQL, rutas y procedimientos. Una ruta invocada en SQL es una función o procedimiento que se invoca desde SQL. Ambas funciones y procedimientos son tipos de instrucciones de SQL que pueden manejar parámetros (valores pasados a una instrucción cuando se invoca esa instrucción). Una <i>función</i> puede recibir parámetros de entrada y devolver un valor basado en la expresión incluida en la instrucción de función. Un <i>procedimiento</i> puede recibir y devolver parámetros de entrada y salida. (Ver el capítulo 13.)
Activadores	Un objeto asociado con una tabla base que define una acción que debe tomarse cuando un evento se produce en relación con esa tabla. La acción que causa la activación del activador (ejecución) puede ser un insertar en, eliminar de, o la actualización de una tabla base. Por ejemplo, la eliminación de una fila en una tabla puede causar la activación de un activador que luego borra los datos de otra tabla. (Ver el capítulo 14.)
Rutina invocada por SQL	Una función o procedimiento que se invoca desde SQL. Una ruta invocada en SQL puede estar en un esquema de objeto o incrustado en un módulo, que también es un objeto de esquema. (Ver el capítulo 13.)
Conjunto de caracteres	Una colección de atributos de caracteres que definen cómo se representan. Un conjunto de caracteres tiene tres atributos: el repertorio, la forma de uso y el cotejo predeterminado. El <i>repertorio</i> determina qué caracteres se expresan (por ejemplo, A, B, C, y así sucesivamente). La <i>forma de uso</i> determina cómo se representan los caracteres como cadenas de hardware y software (por ejemplo, A viene antes que B, B viene antes que C). El <i>cotejo predeterminado</i> determina cómo esas cadenas se comparan unas con otras.
Cotejo	Un conjunto de reglas que controlan cómo las cadenas de caracteres se comparan unas con otras dentro de un repertorio en particular. Esta información se usa para ordenar los caracteres (por ejemplo, A viene antes que B, B viene antes que C). Un cotejo predeterminado se define por cada conjunto de caracteres.
Transliteración	Una operación que mapea caracteres de un conjunto de caracteres a otro conjunto de caracteres. Las transliteraciones incluyen operaciones tales como la traducción de caracteres de mayúsculas a minúsculas o de un alfabeto en otro.
Generador de secuencia	Un mecanismo para generar valores sucesivos de datos numéricos (enteros), uno a la vez. El generador de secuencia mantiene un <i>valor de base actual</i> , que se usa como la base para generar el siguiente valor secuencial.

Tabla 2-2 Tipos de objetos que se definen en cada esquema (*continuación*).

Como se comentó, se analizarán la mayoría de los temas de la tabla con mayor detalle más adelante en el libro. Sin embargo, los últimos tres temas, los cuales están relacionados con los conjuntos de caracteres, se tratan brevemente. Los conjuntos de caracteres, cotejos y transliteraciones respaldados por los RDBMS pueden variar de un producto a otro, y también la implementación de esas características. A lo largo de este libro, todos los ejemplos y proyectos que se darán cuentan con cualquier conjunto de caracteres predeterminados para el producto que utiliza. Si se quiere cambiar ese conjunto de caracteres, ya sea en un nivel predeterminado o en la base de datos o en la tabla nivel, primero debe revisar cuidadosamente la documentación del producto para averiguar qué se respalda y cómo se implementan esas características.

Pregunta al experto

- P:** Se describe un dominio como un objeto definido por el usuario que se basa en un tipo de datos pero puede incluir un valor predeterminado y una restricción. ¿Cómo difiere este tipo de dominio de un dominio como el que se describió en el modelo relacional?
- R:** En muchos casos los dos son lo mismo, y para todos los efectos prácticos se puede pensar en un dominio de SQL como contraparte a un dominio en el modelo relacional. Hay una sutil diferencia, no obstante [un dominio en el modelo relacional es meramente una descripción de los datos que se incluyen en un atributo (columna) asociado con ese dominio en particular]. Por otro lado, un dominio de SQL restringe los datos que se pueden insertar en una columna. Un dominio de SQL hace esto a través del uso de restricciones, las cuales son reglas de validación que son parte del sistema de integridad de los datos. La idea principal a tener en cuenta es que un dominio en el modelo relacional es un concepto lógico, mientras que un dominio de SQL es un concepto físico.
- P:** Cuando se habló acerca de objetos de esquema, se mencionaron las tablas base. ¿Respalda SQL cualquier otro tipo de tablas?
- R:** El estándar SQL:2006 respalda cuatro tipos de tablas: tablas base, tablas transitorias, tablas derivadas y tablas vistas. La tabla base es un tipo de tabla cuyos datos se almacenan en alguna parte. En otras palabras, los datos de SQL se almacenan en una tabla base. Una tabla transitoria es una tabla nombrada que implícitamente se crea durante la evaluación de una expresión de consulta o la ejecución de un activador. Una tabla derivada es aquella que contiene los resultados de una consulta (el conjunto de datos especificados en la consulta). Una tabla vista es otro nombre que se le da a una vista, la cual es una tabla virtual cuya definición se almacena, pero cuyos datos se derivan de tablas base en el momento en que se llama la vista.

¿Qué es una base de datos?

Como habrá notado, en ninguna parte de la estructura del entorno de SQL o de un catálogo se hace mención de una base de datos. La razón es que en ninguna parte del estándar SQL se define el término “base de datos”. De hecho, la única mención de una base de datos, en términos de cómo

podría encajar en la estructura del entorno de SQL, es que pueda considerar a los sitios como la base de datos, aunque esto se ofrece como una sugerencia más que como una definición absoluta. Si bien el estándar utiliza la palabra para referirse a SQL como el lenguaje de una base de datos, en realidad nunca se define el término.

Este enfoque puede estar bien para el estándar, pero en el mundo real puede ser difícil para un RDBMS crear un entorno SQL sin crear algún tipo de componentes que los usuarios puedan señalar y decir: “Sí, es la base de datos.” Y, de hecho, la mayoría de los productos permiten crear, alterar y eliminar objetos que se llaman bases de datos. En SQL Server, por ejemplo, una instancia del software de DBMS puede gestionar cualquier número de bases de datos, siendo cada base de datos una colección lógica de base de datos objetos que el diseñador escoge para administrar juntos. Sybase, MySQL y DB2 de IBM tienen arquitecturas similares. SQL Server proporciona una consola de administración llamada Microsoft SQL Server Management Studio para ver y manipular la base de datos objetos. El panel Object Explorer a lo largo del margen izquierdo proporciona una estructura de directorio jerárquica que incluye una base de datos nodo, con cada base de datos mostrada debajo, y los objetos tales como tablas figuran abajo de cada base de datos. La figura 2-3 muestra el SQL Server Management Studio con la base de datos INVENTARIO expandida hacia abajo de las columnas de la tabla CDS_ARTISTA.

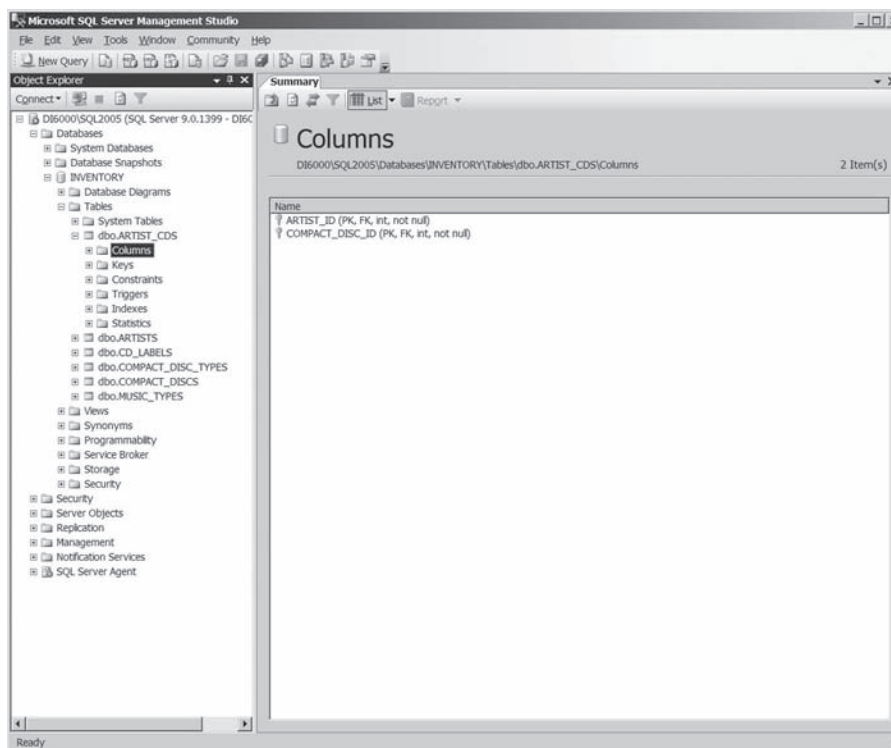


Figura 2-3 Microsoft SQL Server Management Studio con la base de datos INVENTARIO expandida.

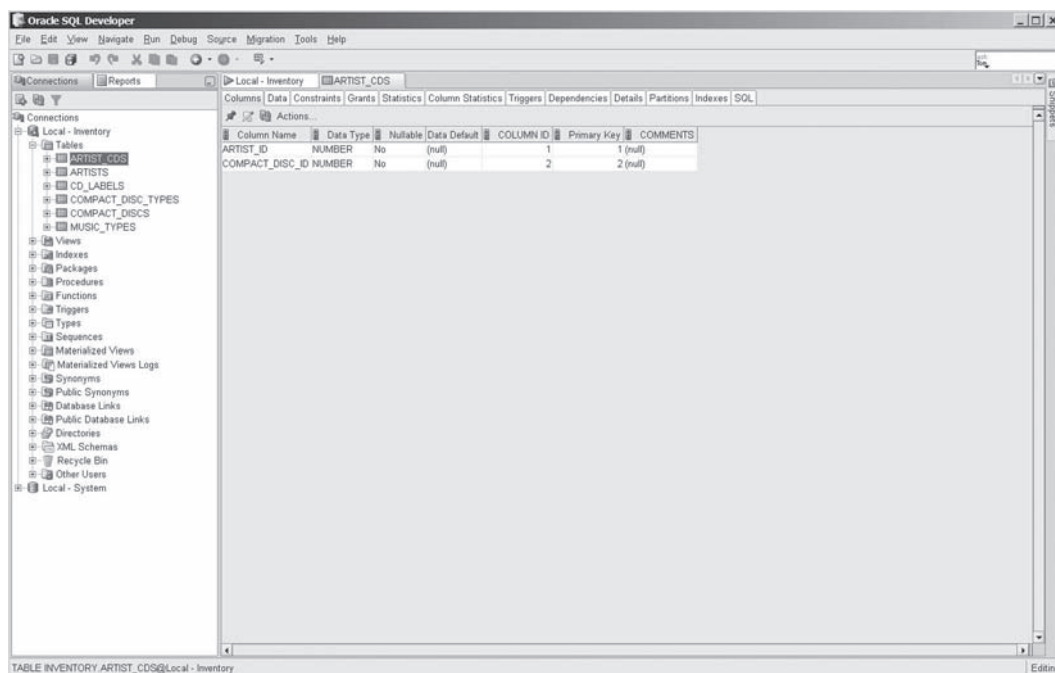


Figura 2-4 Oracle SQL Developer con el esquema INVENTARIO expandido.

El DBMS Oracle tiene una arquitectura diferente. Cada instancia del software de DBMS administra sólo una base de datos. Sin embargo, el usuario de cada base de datos obtiene un esquema distinto para almacenar una base de datos objetos que pertenece a ese usuario. Por lo tanto, en Oracle un esquema es muy similar a lo que SQL Server llama una base de datos. Oracle tiene una herramienta llamada SQL Developer que es funcionalmente similar al SQL Server Management Studio. La figura 2-4 muestra SQL Developer con el esquema INVENTARIO expandido hacia abajo de las columnas en la tabla CDS_ARTISTA. Oracle fue el primer RDBMS comercialmente disponible, y como fue creado mucho antes de que existiera un estándar SQL, no debe ser sorpresa que sea arquitectónicamente diferente.

En el capítulo 1 se expuso que una base de datos es una colección de datos organizados en un formato estructurado que se define por los metadatos que describen esa estructura. En ambos, SQL Server y Oracle, se puede observar cómo se aplica esta definición. Los dos sistemas (y cualquier RDBMS con el que trabaje) reúnen los datos en un formato estructurado y definen esos datos a través del uso de esquemas, los cuales contienen los metadatos. Esta definición también se puede aplicar al estándar SQL y a la construcción de su entorno y a los catálogos. Los datos de SQL se almacenan en un formato organizado dentro de tablas base. Estas tablas base figuran dentro de un esquema, que define esas tablas, y de este modo define los datos. Así que aunque el estándar SQL:2006 no define en realidad el término “base de datos”, no obstante apoya su concepto, al igual que los productos RDBMS que aplican SQL.

Nombrado de objetos en un entorno SQL

Hasta este punto en el libro se ha proporcionado gran cantidad de información conceptual y antecedentes. La razón de esto es que quiero que haya un fundamento básico en SQL antes de empezar a escribir instrucciones SQL. Creo que con esta información estará en mejores condiciones de comprender la lógica detrás del código SQL que creará y la razón para crearlo, y no cabe duda de que está más que listo para comenzar a escribir esas instrucciones.

Sin embargo, antes de empezar a entrar de lleno a SQL, hay un tema más que se necesita cubrir brevemente: los identificadores de objeto. Un *identificador* es un nombre dado a un objeto de SQL. El nombre puede ser de hasta (pero no incluir) 128 caracteres, y debe seguir los convenios definidos. Un identificador se puede asignar a cualquier objeto que se crea con instrucciones SQL, tales como dominios, tablas, columnas, vistas o esquemas. El estándar SQL:2006 define dos tipos de identificadores: identificadores regulares e identificadores delimitados.

Los identificadores regulares son bastante restrictivos y deben seguir convenios específicos:

- Los nombres no se distinguen entre mayúsculas y minúsculas. Por ejemplo, Nombres_Artista es lo mismo que NOMBRES_ARTISTA y nombres_artista.
- Sólo se permiten letras, dígitos y guiones. Por ejemplo, se pueden crear identificadores tales como Primer_Nombre, 1erNombre o PRIMER_NOMBRE. Observe que el guión bajo es el único carácter válido que se usa como separador entre palabras. Los espacios no son aceptables ni tampoco guiones (los guiones se interpretan como operaciones de sustracción).
- No se puede utilizar palabras clave reservadas en SQL.

NOTA

Una palabra clave es una palabra que es parte del léxico de SQL. Hay dos tipos de palabras clave en SQL: reservadas y no reservadas. Como el nombre sugiere, las palabras clave reservadas no se pueden usar para cualquier propósito, ya que están destinadas a utilizarse dentro de una instrucción SQL. Las palabras no reservadas no tienen dichas restricciones. Para un listado completo de palabras clave de SQL, vea el apéndice B.

SQL no distingue mayúsculas y minúsculas, por lo que respecta a los identificadores regulares. Todos los nombres se cambian a mayúsculas cuando se almacenan en SQL, que es la razón por la que 1erNombre y 1ERNOMBRE se leen como valores idénticos. Como ya se mencionó, el caso de falta de distinción a utilizar mayúsculas y minúsculas es el comportamiento predefinido en la mayoría de los RDBMS, y mientras que la predefinición se puede modificar en algunos productos, se recomienda que no se cambie, ya que no sería coherente con el estándar SQL y daría lugar a problemas de compatibilidad en caso de que utilice otros productos para acceder a sus datos.

Los identificadores delimitados no son tan restrictivos como los identificadores regulares, pero aún deben seguir convenios específicos:

- El identificador debe estar incluido en un conjunto de comillas dobles, como el identificador "NombresArtista".
- Las comillas no se almacenan en la base de datos, pero todos los demás caracteres se almacenan como aparecen en la instrucción SQL.

- Los nombres son sensibles a mayúsculas y minúsculas. Por ejemplo, “Nombres_Artista” *no es* lo mismo que “nombres_artista” o “NOMBRES_ARTISTA”, pero “NOMBRES_ARTISTA” *es* lo mismo que NOMBRES_ARTISTA y Nombres_Artista (porque los identificadores regulares se cambian a mayúsculas).
- La mayoría de los caracteres están permitidos, incluyendo espacios.
- Se pueden utilizar palabras clave reservadas a SQL.

Cuando se decide cómo nombrar los objetos SQL, hay una serie de sistemas que se pueden seguir. La primera elección que tendrá que hacer es si desea utilizar los identificadores regulares o delimitados. También decidirá otras cuestiones, tales como utilizar mayúsculas o minúsculas, y el uso del guión bajo, y si los identificadores estarán en singular o plural. Por ejemplo, se puede nombrar una tabla como `TitulosDiscoCompacto`, `titulos_disco_compacto`, `TITULOS_DISCO_COMPACTO`, “`Titulos Disco Compacto`”, o alguna otra forma de ese nombre. La parte importante para recordar es que debe elegir un nombre convencional y adherirse a él a lo largo de la codificación para una base de datos en particular. También debe tomar en cuenta el RDBMS que utilizará. Versiones anteriores del estándar SQL sólo permitían letras mayúsculas en los nombres de objetos (como consecuencia de los sistemas centrales que usaban una colección de caracteres llamados EBCDIC, que en los primeros días sólo contenían letras mayúsculas). Como resultado, muchos de los productos RDBMS más antiguos, incluyendo Oracle y DB2, automáticamente convertían los nombres de minúsculas a mayúsculas. Por lo tanto, se podía crear una tabla llamada `TitulosDiscoCompacto`, pero aparecía en el catálogo como `TITULOSDISCOCOMPACTO`, que obviamente no es un nombre fácil de usar. Por cierto, MySQL sigue la sensibilidad a mayúsculas y minúsculas del sistema operativo, así que en Windows no es sensible a mayúsculas y minúsculas, pero Linux y Unix son sensibles a mayúsculas y minúsculas. Y finalmente, tenga en cuenta la extensión máxima del identificador en el producto RDBMS que utilice. Mientras que SQL Server permite 128 caracteres para los nombres, Oracle admite hasta 30 caracteres (8 para los nombres de base de datos), y muchos otros tienen una extensión máxima que es inferior a 128.

NOTA

Para los ejemplos y proyectos en este libro, se usan identificadores regulares con letras mayúsculas y los guiones bajos se utilizan para separar las palabras (por ejemplo, `TITULOS_DISCO_COMPACTO`). Se hace esto porque dichos identificadores son compatibles con (y forman nombres de objeto fáciles de utilizar) todos los productos RDBMS. Sin embargo, reconozco que a los usuarios de SQL Server les agrada en particular utilizar identificadores con mayúsculas y minúsculas.

Nombres calificados

Todos los identificadores de esquema de objeto se califican por la forma lógica en la que encajan en la estructura jerárquica del entorno SQL. Un nombre completo calificado incluye el nombre del catálogo, el nombre del esquema y el nombre del objeto de esquema, cada uno separado por un punto. Por ejemplo, suponga que tiene una tabla llamada `CD_ARTISTAS`. La tabla está en el esquema `DISCOS_COMPACTOS`, que está en el catálogo `MUSICA`. El nombre completo calificado para esa tabla sería `MUSICA.DISCOS_COMPACTOS.CD_ARTISTAS`.

La forma en que se desempeñan esos convenios de nomenclatura en diferentes productos RDBMS depende de cómo ese producto aplica la estructura del entorno SQL. Por ejemplo, un

nombre completo calificado en SQL Server se basa en el nombre del servidor, nombre de la base de datos, nombre del propietario y nombre del objeto. En este caso, una tabla llamada ARTISTAS puede tener el nombre completo calificado de SERVIDOR01.BD_MUSICA.PBD.ARTISTAS, donde SERVIDOR01 es el nombre del servidor, BD_MUSICA es el nombre de la base de datos, y PBD (que se refiere al propietario de la base de datos) es el nombre del propietario del objeto. Para determinar cómo se manejan los nombres completos calificados para un RDBMS en particular, verifique la documentación del producto.

Creación de un esquema

Ahora que tiene una comprensión fundamental de cómo se utilizan los identificadores para nombrar los objetos SQL, está listo para empezar a escribir instrucciones SQL. Se empezará con la instrucción CREATE SCHEMA, ya que los esquemas están en la parte superior de la jerarquía de SQL, en términos de los objetos que el estándar SQL:2006 le permite crear. (Recuerde, el estándar SQL no proporciona ninguna clase de instrucción como CREATE CATALOG o CREATE DATABASE. Se deja a los proveedores de RDBMS determinar cómo y si se aplica a estos objetos.) Y como ya se mencionó, Oracle crea automáticamente un esquema para cada usuario; así, mientras tiene una instrucción CREATE SCHEMA, es sólo por compatibilidad con el estándar SQL. En la siguiente sección, “Creación de una base de datos”, saldré del modo SQL y analizaré la creación de base de datos, ya que la mayoría de los productos RDBMS respaldan la creación de objetos de base de datos, y es probable que encuentre que desea crear una base de datos con el fin de probar los ejemplos y proyectos en este libro.

El lugar para empezar con cualquier tipo de instrucción SQL es la sintaxis que define las instrucciones. La siguiente sintaxis muestra los componentes básicos de la instrucción CREATE SCHEMA:

```
CREATE SCHEMA <nombre de la cláusula>
[ <conjunto de caracteres o ruta> ]
[ <elementos del esquema> ]
```

NOTA

Los corchetes angulados contienen información que sirve como un marcador de posición para un valor o una cláusula en relación con esa información. Por ejemplo, <nombre de la cláusula> es el marcador de posición para palabras clave y valores relacionados con el nombramiento del esquema. Por otro lado, las llaves significan que la cláusula es opcional. No es necesario especificar un conjunto de caracteres, ruta o el elemento del esquema.

Echemos un vistazo a la sintaxis de la instrucción CREATE SCHEMA pieza por pieza. Las palabras clave CREATE SCHEMA activan la aplicación de SQL con el tipo de instrucción que se ejecuta. Esto continúa con el marcador de posición <nombre de la cláusula>, el cual puede incluir un nombre para el esquema, un identificador de autorización (precedido por la palabra clave AUTHORIZATION), o ambos. Como resultado, el nombre de la cláusula puede tener cualquiera de las siguientes formas:

- <nombre del esquema>
- AUTHORIZATION <identificador de autorización>
- <nombre del esquema> AUTHORIZATION <identificador de autorización>

El valor <identificador de autorización> especifica quién es el propietario del esquema y sus objetos. Si ninguno se especifica, el valor predetermina el del usuario actual. Si no se especifica el valor <nombre del esquema>, se crea un nombre que se base en el identificador de autorización.

La siguiente cláusula, <conjunto de caracteres o ruta>, le permite definir un conjunto de caracteres predeterminados, una ruta predeterminada, o ambos. El nombre del conjunto de caracteres se precede por las palabras clave DEFAULT CHARACTER SET y especifican un conjunto de caracteres predeterminados para el nuevo esquema. La ruta especifica una orden para buscar rutinas invocadas por SQL (procedimientos y funciones) que se crean como parte de la instrucción CREATE SCHEMA. (Las rutinas invocadas por SQL se analizan en el capítulo 13.)

La cláusula <elementos del esquema> se compone de varios tipos de instrucciones de SQL que se pueden incluir en la instrucción CREATE SCHEMA. En su mayor parte, esta cláusula permite crear objetos de esquema tales como tablas, vistas, dominios y activadores. La ventaja de esto es que los objetos se añaden correctamente al esquema cuando se crea, todo en un solo paso.

Ahora que ha visto la sintaxis para la instrucción CREATE SCHEMA, veamos un ejemplo. El siguiente código crea un esquema llamado INVENTARIO. La instrucción también especifica un nombre de identificador de autorización MNGR y un conjunto de caracteres llamado Latino1.

```
CREATE SCHEMA INVENTARIO AUTHORIZATION MNGR
DEFAULT CHARACTER SET Latino1
CREATE TABLE ARTISTAS
( ID_ARTISTA INTEGER, NOMBRE_ARTISTA CHARACTER (20) );
```

Observe que el código de ejemplo incluye la instrucción CREATE TABLE. Éste es uno de los elementos que se puede especificar como parte de la cláusula <elementos del esquema>. Se pueden incluir tantas instrucciones como se quiera. Esta instrucción en particular crea tablas llamadas ARTISTAS que contienen la columna de ID_ARTISTA y la columna NOMBRE_ARTISTA. (Se analizará la instrucción CREATE TABLE con mayor detalle en el capítulo 3.)

Además de definir la instrucción CREATE SCHEMA, SQL:2006 también define la instrucción DROP SCHEMA, como se muestra en la siguiente sintaxis:

```
DROP SCHEMA <nombre del esquema>
CASCADE | RESTRICT
```

La primera línea es bastante sencilla: el esquema nombrado será eliminado del sistema. La segunda línea tiene dos opciones: CASCADE y RESTRICT.

NOTA

El símbolo de barra vertical (|) se puede leer como “o”, lo que significa que se debe utilizar la opción CASCADE o la opción RESTRICT, pero no ambas.

Si se especifica la opción CASCADE, todos los objetos de esquema y los datos de SQL dentro de esos objetos se eliminan del sistema. Si se usa la opción RESTRICT, el esquema se elimina sólo si no existen objetos de esquema. Este método se utiliza como protección contra la eliminación de cualquier objeto que no se desea borrar. Es una forma de comprobar que los objetos que se eliminan son los que se desean borrar antes de que realmente se elimine el esquema.

Ahora veamos un ejemplo de la instrucción DROP SCHEMA. El siguiente código elimina el esquema INVENTARIO:

```
DROP SCHEMA INVENTARIO CASCADE;
```

Observe que se utiliza la opción CASCADE, lo que significa que todos los objetos de esquema y los datos de SQL se eliminarán.

Creación de una base de datos

A pesar del hecho de que el estándar SQL no define qué es una base de datos, y mucho menos proporciona una instrucción para crear cualquier tipo de base de datos objeto, hay una gran posibilidad de que trabaje con un RDBMS que no sólo respalde la creación de una base de datos objeto, sino que también utilice a ese objeto como base para su estructura jerárquica en la gestión de datos objetos. Por consiguiente, se puede encontrar que, con el fin de trabajar a través de los ejemplos y proyectos de este libro, deseará crear una base de datos de prueba para tener un entorno en el cual pueda crear, modificar o eliminar los datos objetos o los datos según sea necesario, sin correr el riesgo de la pérdida de las definiciones de datos o información de la base de datos actual. (Idealmente, se trabaja con un RDBMS que está limpio de instalaciones, sin ningún tipo de base de datos existentes, excepto el sistema preinstalado y base de datos de prueba.)

Si ya ha trabajado con un RDBMS, puede ser que esté familiarizado con la forma en que los objetos de base de datos están organizados dentro de un sistema. Por ejemplo, si echa un vistazo de nuevo a la figura 2-3, puede ver que SQL Server organiza las bases de datos del servidor en una estructura lógica debajo del nodo Bases de datos. Cada nodo Bases de datos (por ejemplo, INVENTARIO) contiene nodos secundarios que representan los diferentes tipos de objetos asociados con esa base de datos en particular. Como se puede ver, la base de datos INVENTARIO actualmente enumera ocho categorías de objetos: diagramas de base de datos, tablas, vistas, sinónimos, programación, agente de servicio, almacenamiento y seguridad. Y debajo de la tabla CDS_ARTISTA, las categorías son columnas, claves, restricciones, desencadenadores, índices y estadísticas. Para una definición de cómo SQL Server define cada uno de esos tipos de objetos, tiene que examinar la documentación del producto, que debe hacer para cualquier RDBMS. Compare y contraste con las categorías de objetos de Oracle mostradas en la figura 2-4.

La mayoría de los productos que respaldan objetos de base de datos también apoyan el lenguaje para crear esos objetos. Por ejemplo, Oracle, MySQL y SQL Server incluyen la instrucción CREATE DATABASE en los lenguajes basados en SQL. Sin embargo, qué parámetros se pueden definir cuando se construyen esas instrucciones, qué permisos se necesitan con el fin de ejecutar esas instrucciones y cómo un sistema implementa los objetos de base de datos varía de un producto a otro. Afortunadamente, la mayoría de los productos utilizan la misma sintaxis básica para crear una base de datos objeto:

```
CREATE DATABASE <nombre de la base de datos>  
<parámetros adicionales>
```

Antes de crear una base de datos en cualquier sistema, asegúrese de leer primero la documentación del producto, y si es apropiado, consulte con el administrador de base de datos para tener la certeza que es seguro agregar una base de datos objeto en el entorno de SQL. Una vez que se crea la base de datos, se pueden crear esquemas, tablas, vistas y otros objetos dentro de esa base de datos, y desde allí llenar las tablas con los datos necesarios.

Pruebe esto 2-1 La creación de una base de datos y un esquema

En “Pruebe esto 1-2: Conexión a una base de datos” se estableció la forma de acceder a un RDBMS. En ese proyecto se usó una aplicación de usuario que permite invocar directamente instrucciones SQL. Se utilizará esa aplicación para este proyecto (y el resto de los proyectos en el libro) para crear una base de datos y un esquema, o cualquiera de esas funciones que el sistema respalda. Una vez que se crea la base de datos, se trabajará dentro del contexto de esa base de datos para futuros ejemplos y proyectos. Si el sistema respalda la creación de esquemas pero no la creación de base de datos, debe trabajar dentro del contexto de ese esquema para los otros proyectos.

Paso a paso

1. Abra la aplicación de cliente que le permita invocar directamente las instrucciones SQL. Si es aplicable, consulte con el administrador de base de datos para asegurarse de que está entrando con las autorizaciones necesarias para crear una base de datos y un esquema. Puede que necesite permisos especiales para crear esos objetos. También verifique si hay algunos parámetros que debe incluir cuando se crea la base de datos (por ejemplo, el tamaño del archivo de registro), restricciones en el nombre que utilizará o restricciones de cualquier otro tipo. Asegúrese de comprobar la documentación del producto antes de seguir adelante.
2. Cree una base de datos llamada INVENTARIO (si el RDBMS respalda esta funcionalidad, en Oracle se crea un nombre de usuario llamado INVENTARIO, que implícitamente crea un esquema con el mismo nombre). Dependiendo del producto que se use, se ejecutará una instrucción que sea similar a la siguiente:

```
CREATE DATABASE INVENTARIO;
```

Si se requiere incluir parámetros adicionales en la instrucción, lo más probable es que se incluyan en las líneas siguientes a la cláusula CREATE DATABASE. Una vez que se ejecute la instrucción, debe recibir algún tipo de mensaje que indique que la instrucción se ejecutó con éxito.

3. Conexión a la nueva base de datos. El método para hacer esto varía de un producto a otro. En Oracle se puede conectar a la base de datos introduciendo la información de inicio de sesión en cualquiera de las muchas herramientas tales como SQL*Plus, iSQL*Plus y SQL Developer. En SQL Server, es simplemente cuestión de seleccionar la base de datos apropiada de la lista de conexión desplegable de las bases de datos en la barra de herramientas del SQL Server Management Studio, o se puede ejecutar la siguiente instrucción (MySQL utiliza la misma sintaxis):

```
USE Inventario
```

4. Cree un esquema llamado CD_INVENTARIO (si el RDBMS respalda esta funcionalidad). Cree un esquema bajo su actual identificador de autorización. No incluya ninguno de los elementos del esquema en este momento. En la mayoría de los casos, se ejecutará una instrucción que sea similar a la siguiente:

```
CREATE SCHEMA CD_INVENTARIO;
```

(continúa)

Resumen de Pruebe esto

Los ejercicios paso a paso de este tipo pueden ser complicados porque dependen de la forma en que los productos RDBMS implementan diversas características. Como resultado, debe confiar en gran medida en la documentación del producto (que debe utilizar de todos modos) y, en su caso, a los administradores de base de datos. Sin embargo, ahora que consiguió pasar por este ejercicio y se creó la base de datos necesaria y/o el esquema del entorno, debe estar listo para continuar con los ejemplos y proyectos del resto del libro. Ya que se han sentado las bases, estará listo para crear, alterar, suprimir los objetos de datos, e insertar, modificar y borrar los datos almacenados en esos objetos.

Autoexamen Capítulo 2

1. ¿Cuáles son las diferencias entre un agente SQL y una implementación de SQL?
2. ¿Qué componente del entorno SQL representa a un usuario o rol que concede privilegios específicos para acceder a los objetos y datos?
 - A Catálogo
 - B Identificador de autorización
 - C Módulo de cliente SQL
 - D Agente SQL
3. Un(a) _____ es una colección de esquemas que forman un nombre dentro del entorno SQL.
4. ¿Qué es un esquema?
5. ¿Qué instrucción se utiliza para agregar un esquema en el entorno SQL?
 - A ADD SCHEMA
 - B INSERT SCHEMA
 - C CREATE SCHEMA
6. ¿Cuál es el nombre del esquema que contiene las definiciones para los objetos de esquema en un catálogo?
7. ¿Cuáles son los 11 tipos de objetos de esquema que pueden estar contenidos en un esquema?
8. ¿Qué es una vista?

- 9.** ¿Cuáles objetos de esquema proporcionan la unidad básica de gestión de datos en el entorno SQL?
 - A** Vistas
 - B** Dominios
 - C** Tablas base
 - D** Conjunto de caracteres
- 10.** ¿Cómo define el estándar SQL:2006 a una base de datos?
- 11.** Un(a) _____ es el nombre dado a un objeto SQL.
- 12.** ¿Cómo se distingue un identificador regular de un identificador delimitado en una instrucción SQL?
- 13.** ¿Qué tipo de identificador permite que se utilicen espacios como parte del nombre de un objeto?
- 14.** El entorno SQL incluye un catálogo llamado INVENTARIO. En ese catálogo está el esquema llamado DISCOS_COMPACTOS, y en ese esquema se encuentra una tabla denominada ARTISTAS. ¿Cuál es el nombre cualificado de esa tabla?
- 15.** ¿Cuáles son las tres formas que puede tomar el componente <cláusula de nombre> de una instrucción CREATE SCHEMA?
- 16.** ¿Cuáles son las diferencias entre la opción CASCADE y la opción RESTRICT en la instrucción DROP SCHEMA?
- 17.** Dentro de la jerarquía del entorno SQL, ¿cómo está relacionado un dominio con un catálogo?
- 18.** ¿Qué tipo de identificador permite utilizar una palabra clave reservada?

Capítulo 3

Creación y modificación de tablas

Habilidades y conceptos clave

- Creación de tablas en SQL
 - Especificación de los tipos de datos en una columna
 - Creación de tipos definidos por el usuario
 - Especificación de los valores predeterminados en una columna
 - Modificación de tablas en SQL
 - Eliminación de tablas en SQL
-

En el entorno SQL, las tablas son la unidad básica de gestión de datos. La mayoría de la programación que se hace en SQL se relaciona directa o indirectamente con esas tablas. Como resultado, antes de insertar la información en la base de datos o modificar esa información, las tablas apropiadas deben haberse creado o debe crearlas. El estándar SQL:2006 proporciona tres instrucciones que permiten definir, cambiar o eliminar las definiciones de las tablas en el entorno SQL. Se puede utilizar la instrucción `CREATE TABLE` para añadir una tabla, la instrucción `ALTER TABLE` para modificar esa definición, o la instrucción `DROP TABLE` para eliminar la tabla y sus datos de la base de datos. De esas tres instrucciones, la instrucción `CREATE TABLE` tiene la sintaxis más compleja. No sólo es esto por los distintos tipos de tablas respaldadas por SQL, sino también porque la definición de la tabla puede incluir muchos elementos. Sin embargo, a pesar de esas complejidades, la creación de una tabla es un proceso bastante sencillo, una vez que se comprende la sintaxis básica.

Creación de tablas en SQL

Como puede que recuerde del capítulo 2, SQL respalda tres tipos de tablas: tablas base, tablas derivadas y tablas vistas. La mayoría de las tablas base son objetos de esquema que tienen los datos de SQL. Las tablas derivadas son los resultados que se observan cuando se solicitan (consultan) datos de una base de datos. Las tablas vistas son otro nombre para las vistas, que se analizarán en el capítulo 5. Se puede pensar en las tablas vistas como un tipo de nombre derivado de una tabla, con una definición de la vista almacenado en el esquema.

En este capítulo se trabajará con tablas base. De hecho, la mayoría de lo que estará trabajando directamente a lo largo de este libro (como también a lo largo de su carrera como programador) son tablas base; sin embargo, no todas las tablas base son lo mismo. Algunas son persistentes (permanentes) y otras son temporales. Algunas están en objetos de esquema y otras están contenidas en módulos. Todos los módulos de tablas base son también tablas temporales. SQL respalda cuatro tipos de tablas base:

- **Tablas base persistentes** Un objeto de esquema nombrado definido por la definición de una tabla en la instrucción `CREATE TABLE`. Las tablas base persistentes tienen los datos de SQL

que se almacenan en la base de datos. Éste es el tipo más común de tabla base y es a menudo a lo que se refiere la gente cuando menciona tablas base o tablas. Una tabla base persistente existe desde que la definición de tabla existe, y se puede llamar desde cualquier sesión de SQL.

- **Tablas temporales globales** Un objeto de esquema nombrado definido por una definición de tabla en la instrucción `CREATE GLOBAL TEMPORARY TABLE`. Aunque la definición de la tabla es parte del esquema, la tabla actual existe sólo cuando se hace referencia dentro del contexto de la sesión SQL en la cual se creó. Cuando la sesión termina, la tabla ya no existe. No se puede acceder a una tabla temporal global creada en una sesión desde otra sesión de SQL. Los contenidos son distintos en cada sesión de SQL.
- **Tablas temporales locales creadas** Un objeto de esquema nombrado definido por una definición de tabla en la instrucción `CREATE LOCAL TEMPORARY TABLE`. Al igual que una tabla temporal global, sólo se puede hacer referencia a una tabla temporal local creada dentro del contexto de la sesión de SQL en la cual se creó, y no se puede acceder desde otra sesión de SQL. Sin embargo, se puede acceder a una tabla global desde cualquier lugar dentro de una sesión asociada de SQL, mientras que en una tabla temporal local sólo se podrá acceder dentro del módulo asociado. Los contenidos son distintos dentro de ese módulo.
- **Tablas temporales locales declaradas** Una tabla declarada como parte de un procedimiento en un módulo. La definición de la tabla no se incluye en el esquema y no existe hasta que ese procedimiento se ejecuta. Al igual que otras tablas temporales, sólo se hace referencia a una tabla temporal local declarada dentro del contexto de la sesión SQL en la cual se creó.

NOTA

Una sesión SQL se refiere a la conexión entre un usuario y un agente SQL. Durante esta conexión, una secuencia de instrucciones SQL consecutivas se invocan por ese usuario y luego se ejecutan. Un módulo es un objeto que contiene instrucciones SQL, rutas o procedimientos. Los módulos se analizarán en el capítulo 13 y en el capítulo 17.

Como puede ver, se puede utilizar una forma de la instrucción `CREATE TABLE` para crear todos los tipos de tablas, excepto las tablas temporales locales declaradas. A lo largo del resto del capítulo, se analizarán en primer lugar las tablas temporales locales persistentes, aunque se tocará el tema de las tablas temporales en los capítulos subsecuentes. Mientras tanto, veamos la sintaxis en la instrucción `CREATE TABLE`:

```
CREATE [ {GLOBAL | LOCAL} TEMPORARY ] TABLE <nombre de la tabla>
( <elemento de la tabla> [ {, <elemento de la tabla> }... ] )
[ ON COMMIT { PRESERVE | DELETE } ROWS ]
```

NOTA

Las llaves se usan para agrupar elementos. Por ejemplo, en la primera línea de la sintaxis, las palabras clave `GLOBAL | LOCAL` se agrupan juntas. Los corchetes indican que primero debe decidir cómo manejar los contenidos dentro de los corchetes y luego determinar cómo encajan en la cláusula. En la primera línea, debe utilizar ya sea `GLOBAL` o `LOCAL` junto con `TEMPORARY`. Sin embargo, toda la cláusula es opcional. Los tres puntos (en la segunda línea) indican que puede repetir la cláusula tantas veces como sea necesario. En este caso, puede añadir tantas cláusulas `<elemento de la tabla>` como la definición requiera.

La sintaxis que se muestra proporciona sólo los fundamentos de la instrucción CREATE TABLE, que es en realidad mucho más compleja. (La sintaxis y sus explicaciones toman alrededor de 38 páginas del estándar SQL:2006.) Aun así, la sintaxis proporcionada aquí es un fundamento suficiente para crear la mayoría de las tablas que es probable que utilice.

En la primera línea de la sintaxis se designa si la tabla es temporal y se proporciona un nombre a la tabla; por lo tanto, tiene tres opciones:

- CREATE TABLE <nombre de la tabla>
- CREATE GLOBAL TEMPORARY TABLE <nombre de la tabla>
- CREATE LOCAL TEMPORARY TABLE <nombre de la tabla>

Dependiendo del RDBMS en que trabaje, puede que tenga que calificar el nombre de la tabla incluyendo un nombre de esquema, identificador autorizado, o el nombre de la base de datos (por ejemplo, INVENTARIO.ARTISTAS).

NOTA

Hay una serie de variaciones específicas en la aplicación con respecto a las tablas temporales que vale la pena mencionar. Oracle (desde 11g) no tiene una opción LOCAL para crear una tabla temporal; los datos en una tabla temporal se privan de la sesión expresada. IBM DB2 UDB desde 9.1 utiliza el comando, DECLARE GLOBAL TEMPORARY TABLE, para la creación de una tabla temporal global; no parece haber ninguna designación para crear/declarar una tabla temporal local. En SQL Server (desde 2008), las tablas temporales se crean con el comando típico CREATE TABLE, pero los nombres de las tablas temporales locales llevan un prefijo con un signo numérico único (#nombre_tabla), y los nombres de las tablas temporales globales llevan un prefijo con un signo numérico doble (##nombre_tabla).

La segunda línea de la sintaxis permite especificar las partes que componen la tabla, tales como columnas. (Se retomará esto en un momento.) La tercera línea de la sintaxis aplica sólo si se crea una tabla temporal. La cláusula permite especificar si la tabla se debe vaciar o no cuando la instrucción COMMIT se ejecuta. La instrucción COMMIT se usa en una transacción al hacer cambios en la base de datos. Se analizarán las transacciones en el capítulo 16.

Se puede pensar de las cláusulas <elemento de la tabla> como el punto central de la instrucción CREATE TABLE. Es aquí donde se definen las columnas, limitaciones y otros elementos específicos de la tabla que se crea. Se pueden definir una o más cláusulas <elemento de la tabla>. Si se define más de una, se deben separar por comas. De los elementos que puede crear, nos centraremos principalmente en las columnas (en este capítulo) y las restricciones (en el capítulo 4). Echemos un vistazo más de cerca a la sintaxis que se utiliza para definir una columna:

```
<nombre de columna> { <tipo de datos> | <dominio> }  
[ <cláusula predeterminada> ] [ <restricción de columna> ] [ COLLATE  
<nombre de cotejo> ]
```

En la primera línea de la sintaxis debe proporcionar el nombre de la columna y declarar el tipo de datos o el dominio definido por el usuario. Se analizarán los tipos de datos en la sección “Especificación de los tipos de datos en una columna”, más adelante en este capítulo, y los dominios en el capítulo 4.

En la segunda línea de la sintaxis se tiene la opción de proporcionar un valor predeterminado (ver la sección “Especificación de los valores predeterminados en una columna”), las restricciones de la columna (ver el capítulo 4) o el cotejo (ver el capítulo 2).

En su forma más básica, la instrucción CREATE TABLE podría verse como la siguiente instrucción:

```
CREATE TABLE ARTISTAS
( ID_ARTISTA      INTEGER,
  NOMBRE_ARTISTA  CHARACTER (60) );
```

En esta instrucción se creó una tabla llamada ARTISTAS, una columna llamada ID_ARTISTA y una columna llamada NOMBRE_ARTISTA. La columna ID_ARTISTA se asocia con el tipo de dato INTEGER, y la columna NOMBREA_ARTISTA se asocia con el tipo de datos CHARACTER. Observe que las definiciones de las dos columnas se separan por una coma. También observe que se colocaron las definiciones de las dos columnas en líneas separadas y se alinearon los tipos de datos agregando espacios extra (todo esto es para mejorar la legibilidad, pero aparte de eso es innecesario) (cuando se procesan las instrucciones SQL, los espacios extra y las líneas nuevas simplemente se ignoran). Si se ejecuta la instrucción CREATE TABLE, la tabla será similar a la mostrada en la figura 3-1.

NOTA

Las filas de datos mostradas no estarán en una tabla hasta que se añadan esos datos. Las filas se muestran simplemente para darle una idea del tipo de tabla que esta instrucción podría crear.

Antes de ir más lejos con la discusión de la creación de una tabla, echemos un vistazo más de cerca a los tipos de datos, que juegan un rol integral en la definición de cualquier columna.

ID_ARTISTA: INTEGER	NOMBRE_ARTISTA: CHARACTER(60)
10001	Jennifer Warnes
10002	Joni Mitchell
10003	William Ackerman
10004	Kitaro
10005	Bing Crosby
10006	Patsy Cline
10007	Jose Carreras
10008	Placido Domingo
10009	Luciano Pavarotti

Figura 3-1 Las columnas ID_ARTISTA y NOMBRE_ARTISTA de la tabla ARTISTAS.

Pregunta al experto

P: Cuando se analizaron los diversos tipos de tablas respaldadas por SQL, se habló brevemente de las tablas temporales. ¿Cuál es el propósito de las tablas temporales?

R: Las tablas temporales proporcionan una forma de almacenar resultados temporales dentro del contexto de su sesión. Puede encontrar que necesita un lugar para almacenar los datos con el fin de tomar un cierto curso de acción. Puede crear explícitamente una tabla base persistente, almacenar datos en ella y después eliminar la tabla cuando termine, pero la tabla temporal permite hacer lo mismo sin tener que destruir explícitamente la tabla cada vez que la use. En otras palabras, una tabla temporal es una herramienta útil cuando se requiere almacenar datos por sólo un periodo específico. Por ejemplo, suponga que tiene una aplicación que permita generar un informe trimestral basado en el inventario al final del periodo que se examina. La aplicación necesitará reunir la información en una colección significativa para generar el reporte. Sin embargo, una vez que el reporte se genera, la aplicación ya no necesita almacenar esa información; por lo tanto, la tabla se puede eliminar. Una de las ventajas de utilizar una tabla temporal es que, debido a que es única en una sesión, la tabla no interactúa con otros usuarios o sesiones. Como resultado, el RDBMS no tiene que tomar medidas especiales para bloquear los datos para evitar que otros usuarios apliquen actualizaciones conflictivas a las tablas temporales, y evitar el bloqueo puede resultar en un mejor rendimiento.

Especificación de los tipos de datos en una columna

Cada vez que se define una columna en la instrucción CREATE TABLE, al menos debe proporcionar un nombre para la columna y un tipo de dato asociado o un dominio. El tipo de datos o dominio (que se analiza en el capítulo 4) limita los valores que pueden introducirse en esa columna. Por ejemplo, algunos tipos de datos limitan los valores de una columna a números, mientras otros tipos de datos permiten que se introduzca cualquier carácter. SQL respalda tres formas de tipos de datos:

- **Predefinido** Los tipos de datos predefinidos son los más comunes. Cada tipo de datos predefinido es un elemento nombrado (utilizando una palabra clave en SQL) que limita valores a las restricciones definidas por esa base de datos. SQL incluye cinco formas de tipos de datos predefinidos: cadena, numérico, fecha y hora, intervalo y booleano.
- **Construido** Los tipos de datos contruidos también se denominan elementos, pero tienden a ser más complejos que los tipos de datos predefinidos ya que pueden contener múltiples valores. Los tipos contruidos permiten *construir* estructuras más complicadas que la mayoría de los tipos de datos tradicionales. Una discusión detallada de estos tipos está fuera del alcance de este libro, pero se mencionan para que sepa que existen.

- **Definido por el usuario** Los tipos de datos definidos por el usuario se basan en los tipos predefinidos o definiciones de atributos, y se agregan como objetos de esquema al entorno SQL. SQL respalda dos formas de tipos de datos definidos por el usuario: distinto y estructurado. El tipo distinto se basa en el tipo de dato predefinido, y el tipo estructurado se basa en las definiciones de atributo. Se analizarán los tipos definidos por el usuario en la sección “Creación de tipos definidos por el usuario”, más adelante en este capítulo.

A pesar de que todas las implementaciones de SQL soportan los tipos de datos, qué tipos de datos soportan varían de un producto a otro. Sin embargo, como un programador principiante de SQL, encontrará que la mayoría de las aplicaciones respaldan la forma básica (más tradicional) de tipos de datos, que son los que se usarán en los ejemplos y ejercicios a lo largo del libro. Estos tipos de datos más tradicionales, algunas veces conocidos como tipos primitivos, son parte de los tipos de datos predefinidos de SQL, que se describen en la sección siguiente. No trate de memorizar cada uno de estos tipos, pero comience a familiarizarse con las diferencias entre ellos. Encontrará que, cuando empiece a utilizar tipos de datos específicos, se sentirá más cómodo con éstos. Mientras tanto, remítase a las secciones siguientes tanto como sea necesario siempre que trabaje con definición de tablas o datos SQL.

Tipos de datos de cadena

Los tipos de datos de cadena se componen por tipos que permiten valores basados en los conjuntos de caracteres o en bits de datos. Los valores permitidos por los tipos en cadena pueden ser de longitud fija o variable, dependiendo del tipo específico. SQL define cuatro formas de tipos de datos en cadena:

- **Cadenas de caracteres** Los valores permitidos se deben extraer de un conjunto de caracteres, ya sea de un conjunto predeterminado o de un conjunto definido en el momento que la columna se define. Los tipos de datos en la cadena de caracteres incluyen CHARACTER, CHARACTER VARYING y CHARACTER LARGE OBJECT.
- **Cadenas de carácter nacional** Los valores permitidos son similares a los de las cadenas de caracteres, salvo que el conjunto de caracteres asociados con estos tipos de datos se definen por la aplicación. Como resultado, cuando una cadena de carácter nacional se especifica, los valores asociados con ese tipo de datos deben basarse en el conjunto de caracteres especificado por el sistema de gestión de base de datos relacional (RDBMS) para las cadenas de carácter nacional. Éstas son útiles para almacenar cadenas de caracteres en varios lenguajes en la misma base de datos. Los tipos de datos en la cadena de carácter nacional incluyen NATIONAL CHARACTER, NATIONAL CHARACTER VARYING y NATIONAL CHARACTER LARGE OBJECT.
- **Cadenas de bits** Los valores permitidos se basan en bits de datos (dígitos binarios), en lugar de conjuntos de caracteres o cotejos, lo que significa que estos tipos de datos permiten sólo valores de 0 o 1. SQL respalda dos formas de tipos de datos de cadena de bits: BIT y BIT VARYING.
- **Cadenas binarias** Los valores permitidos son similares a las cadenas de bits, excepto que se basan en bytes (denominado como *octetos* en SQL:2006), en lugar de bits. Como resultado, ningún conjunto de caracteres o cotejos se relacionan con ellas. (Un byte es igual a 8 bis, la

razón por la cual el estándar SQL utiliza el término octeto.) SQL sólo respalda un tipo de datos en una cadena binaria: **BINARY LARGE OBJECT**. Este tipo es útil para almacenar datos binarios puros tales como clips de sonido o imágenes en la base de datos.

Ahora que tiene una descripción general de las formas de tipos de datos en cadena, echemos un vistazo más de cerca a cada uno. La tabla 3-1 describe cada uno de estos tipos de datos y proporciona un ejemplo de una definición de columna que utilice el tipo específico.

Tipo de dato	Descripción/ejemplo
CHARACTER	Especifica el número exacto de caracteres (que debe ser de un conjunto de caracteres) que se almacenará por cada valor. Por ejemplo, si se define el número de caracteres como 10, pero el valor contiene sólo seis caracteres, los cuatro caracteres restantes serán espacios. El tipo de dato puede abreviarse como CHAR . Ejemplo: <code>NOMBRE_ARTISTA CHAR(60)</code>
CHARACTER VARYING	Especifica el mayor número de caracteres (que debe ser de un conjunto de caracteres) que se incluyen en un valor. El número de caracteres almacenados es exactamente el mismo número que el valor introducido; por lo tanto, no se agregan espacios al valor. El tipo de dato puede abreviarse como CHAR VARYING o VARCHAR . Ejemplo: <code>NOMBRE_ARTISTA VARCHAR(60)</code>
CHARACTER LARGE OBJECT	Almacena grandes grupos de caracteres, hasta la cantidad especificada. El número de caracteres almacenados es exactamente el mismo número que el valor introducido; por lo tanto, no se agregan espacios al valor. El tipo de dato puede abreviarse como CLOB . Ejemplo: <code>BIO_ARTISTA CLOB(200K)</code>
NATIONAL CHARACTER	Funciona igual que el tipo de dato CHARACTER , excepto que se basa en una aplicación definida de un conjunto de caracteres. El tipo de dato puede abreviarse como NATIONAL CHAR y NCHAR . Ejemplo: <code>NOMBRE_ARTISTA NCHAR(60)</code>
NATIONAL CHARACTER VARYING	Funciona igual que el tipo de dato CHARACTER VARYING , excepto que se basa en una aplicación definida de un conjunto de caracteres. El tipo de dato puede abreviarse como NATIONAL CHAR VARYING o NCHAR VARYING . Ejemplo: <code>NOMBRE_ARTISTA NCHAR VARYING(60)</code>
NATIONAL CHARACTER LARGE OBJECT	Funciona igual que el tipo de dato CHARACTER LARGE OBJECT , excepto que se basa en una aplicación definida de un conjunto de caracteres. El tipo de dato puede abreviarse como NCHAR LARGE OBJECT o NCLOB . Ejemplo: <code>BIO_ARTISTA NCLOB(200K)</code>
BIT	Especifica el número exacto de bits que pueden almacenarse para cada carácter. Por ejemplo, si se define el número de bits como 2, pero el valor contiene sólo 1 bit, el bit restante será un espacio. Si el número de bits no se especifica, 1 bit se almacena. Ejemplo: <code>EN_EXISTENCIA BIT</code>

Tabla 3-1 Tipos de datos en cadena con ejemplos de definiciones de columna.

Tipo de dato	Descripción/ejemplo
BIT VARYING	Especifica el mayor número de bits que pueden incluirse en un valor. El número de bits almacenados es exactamente el mismo número que el valor introducido; por lo tanto, no se agregan espacios al valor. Ejemplo: <code>EN_EXISTENCIA BIT VARYING (2)</code>
BINARY LARGE OBJECT	Almacena grandes grupos de bytes hasta la cantidad especificada. El número de bytes almacenados es exactamente el mismo número que el valor introducido; por lo tanto, no se agregan espacios al valor. El tipo de dato también puede remitirse como BLOB. Ejemplo: <code>IMG_ARTISTA BLOB (1M)</code>
XML	El lenguaje de marcado extensible (XML) es un lenguaje de marcado para fines generales utilizado para describir documentos en un formato que es conveniente para la visualización de páginas web y para intercambiar datos entre diferentes partes. Las especificaciones para almacenar datos XML en bases de datos SQL se añaden al estándar SQL en SQL:2003 y se abordan en el capítulo 18. Ejemplo: <code>BIO_ARTISTA XML(DOCUMENT(UNTYPED))</code>

Tabla 3-1 Tipos de datos en cadena con un ejemplo de definiciones de columna (*continuación*).

Tipos de datos numéricos

Como probablemente podrá suponer por el nombre, los valores especificados por los tipos de datos numéricos son números. Todos los tipos de datos numéricos tienen una precisión, y algunos tienen una escala. La *precisión* se refiere al número de dígitos (dentro de un valor numérico específico) que se pueden almacenar. La *escala* se refiere al número de dígitos en la parte fraccional de ese valor (los dígitos a la derecha del punto decimal). Por ejemplo, el número 435.27 tiene una precisión de 5 y una escala de 2. La escala no puede ser un número negativo o ser más larga que la precisión. Una escala de 0 indica que el número es un número entero y no contiene ningún componente fraccional. SQL define dos formas de tipos de datos numéricos:

- **Numéricos exactos** Los valores permitidos tienen precisión y escala, que, para algunos tipos de datos numéricos, se definen por la implementación. Los tipos de datos numéricos exactos incluyen NUMERIC, DECIMAL, INTEGER y SMALLINT.
- **Numéricos aproximados** Los valores permitidos tienen precisión pero no escala. Como resultado, el punto decimal puede flotar. Un número de *punto flotante* es aquel que contiene un punto decimal, pero el punto decimal se puede localizar en cualquier lugar dentro del número, por lo que se dice que un numérico aproximado no tiene escala. Los tipos de datos numéricos aproximados incluyen REAL, DOUBLE PRECISION y FLOAT.

La tabla 3-2 describe cada uno de los tipos de datos numéricos y proporciona un ejemplo de una definición de columna que utiliza un tipo específico.

Tipo de dato	Descripción/ejemplo
NUMERIC	Especifica la precisión y la escala de un valor numérico. Se puede especificar sólo la precisión y utilizar la escala definida (predeterminada) por la aplicación, o se puede especificar la precisión y la escala. Si no se especifica ni la precisión ni la escala, la aplicación proporcionará los valores predeterminados para ambas. Ejemplo: <code>TASA_ARTISTAS NUMERIC (5, 2)</code>
DECIMAL	Especifica valores similares al del tipo de datos NUMERIC. Sin embargo, si la precisión definida por la aplicación es superior a la precisión especificada, los valores con la precisión superior se aceptarán, pero la escala siempre será la que se especifique. Ejemplo: <code>REGALIAS_ARTISTAS DECIMAL (5, 2)</code>
INTEGER	Especifica un valor con una precisión definida por la aplicación y una escala de 0, lo que significa que sólo se aceptan enteros y no se especifica ningún parámetro con este tipo de datos. El tipo de dato puede abreviarse como INT. Ejemplo: <code>ID_ARTISTA INT</code>
SMALLINT	Especifica un valor similar al del tipo de datos INTEGER. Sin embargo, la precisión definida por la aplicación debe ser menor que la precisión de INTEGER. Ejemplo: <code>ID_ARTISTA SMALLINT</code>
FLOAT	Especifica la precisión de un valor numérico, pero no la escala. Ejemplo: <code>REGALIAS_ARTISTAS FLOAT (6)</code>
REAL	Especifica un valor con una precisión definida por la aplicación, pero sin una escala. La precisión debe ser menor a la precisión definida por un tipo de datos DOUBLE PRECISION. Ejemplo: <code>REGALIAS_ARTISTAS REAL</code>
DOUBLE PRECISION	Especifica un valor con una precisión definida por la aplicación, pero sin una escala. La precisión debe ser superior a la precisión definida por el tipo de datos REAL. La implicación es que el valor de la precisión debe ser doble que en el tipo de datos REAL, pero cada implementación define <i>doble</i> de diferentes maneras. Ejemplo: <code>REGALIAS_ARTISTAS DOUBLE PRECISION</code>

Tabla 3-2 Tipos de datos numéricos con un ejemplo de definiciones de columna.

Tipos de datos de fecha y hora

Como su nombre lo indica, los tipos de datos de fecha y hora se refieren a las formas de seguimiento de fechas y horas. SQL define tres formas de tipos de datos de fecha y hora (DATE, TIME y TIMESTAMP) y las variaciones en estos tipos. Estas variaciones están relacionadas con el Tiempo Universal Coordinado (UTC), que solía llamarse el Tiempo Medio de Greenwich (GMT), y las distintas zonas horarias. La tabla 3-3 describe cada uno de los tipos de datos de fecha y hora en el estándar SQL:2006 y proporciona un ejemplo de una definición de columna que utiliza un tipo específico.

Tipo de dato	Descripción/ejemplo
DATE	Especifica el año, mes y día del valor de una fecha. El año tiene cuatro dígitos y respalda valores desde 0001 hasta 9999; el mes tiene dos dígitos y respalda valores desde 01 hasta 12, y el día tiene dos dígitos y respalda valores desde 01 hasta 31. Ejemplo: <code>FECHA_CONTRATACION DATE</code>
TIME	Especifica la hora, minuto y segundo del valor de una hora. La hora tiene dos dígitos y respalda valores desde 00 hasta 23; el minuto tiene dos dígitos y respalda valores desde 00 hasta 59, y el segundo tiene, al menos, dos dígitos y respalda valores desde 00 hasta 61.999 (para dar cabida al salto de segundos). El tipo de datos no incluye dígitos fraccionados a menos que se especifiquen. Por ejemplo, <code>TIME(3)</code> daría tres dígitos fraccionales. El tipo de dato también se puede denominar como <code>TIME WITHOUT TIME ZONE</code> . Ejemplo: <code>HORA_EVENTO TIME(2)</code>
TIME WITH TIME ZONE	Especifica la misma información que el tipo de datos <code>TIME</code> salvo que el valor también incluye información específica del UTC y las zonas horarias. Los valores agregados al tipo de datos tienen un rango desde -11:59 a +12:00. Ejemplo: <code>HORA_EVENTO TIME WITH TIME ZONE(2)</code>
TIMESTAMP	Combina los valores de <code>TIME</code> y <code>DATE</code> . La única diferencia es que con el tipo de datos <code>TIME</code> , el número predeterminado de dígitos fraccionarios es 0, pero con el tipo de datos <code>TIMESTAMP</code> , el número predeterminado es 6. Se puede especificar un número diferente de dígitos fraccionarios incluyendo un parámetro, tal como <code>TIMESTAMP(4)</code> . El tipo de datos se puede también denominar como <code>TIMESTAMP WITHOUT TIME ZONE</code> . Ejemplo: <code>FECHA_COMPRA TIMESTAMP(3)</code>
TIMESTAMP WITH TIME ZONE	Especifica la misma información que el tipo de datos <code>TIMESTAMP</code> , salvo que el valor también incluye información específica del UTC y las zonas horarias. Los valores añadidos al tipo de datos tienen un rango desde -11:59 a +12:00. Ejemplo: <code>FECHA_COMPRA TIMESTAMP WITH TIME ZONE(2)</code>

Tabla 3-3 Tipos de datos de fecha y hora con un ejemplo de definiciones de columna.

Encontrará variaciones de implementación considerables para los tipos de datos de fecha y hora, ya que en un principio las bases de datos no las tenían en absoluto. Esto parece extraño hasta que se da cuenta que se construyeron por informáticos que no sabían cuántas aplicaciones comerciales se basan en fechas y horas. Conforme los usuarios de negocios solicitaron las capacidades de fecha y hora, los proveedores comerciales de RDBMS de hoy en día (mucho antes de que hubiera un estándar SQL) se apresuraron a ofrecer las nuevas características y, por lo tanto, se implementaron en formas muy diferentes. Por ejemplo, el tipo de datos `DATE` de Oracle siempre incluye un componente de hora, y SQL Server utiliza un tipo de datos `TIMESTAMP` para un fin completamente diferente, con un tipo de datos llamado `DATETIME` que funciona como el tipo de datos `TIMESTAMP` de SQL:2006. Por lo tanto, como siempre, consulte la documentación del proveedor.

Tipo de datos de intervalo

El tipo de datos de intervalo está estrechamente relacionado con los tipos de datos de fecha y hora. El valor de un tipo de datos de intervalo representa la diferencia entre dos valores de fecha y hora. SQL respalda dos tipos de intervalos:

- **Intervalos de año-mes** El tipo de datos de intervalo especifica intervalos entre años, meses, o ambos. Se pueden utilizar sólo los campos de AÑO y MES en un intervalo de año-mes.
- **Intervalos de día-hora** El tipo de datos de intervalo especifica intervalos entre cualquiera de los siguientes valores: días, horas, minutos o segundos. Se pueden utilizar sólo los campos de DAY, HOUR, MINUTE y SECOND en un intervalo de día-tiempo.

No se puede mezclar un tipo de intervalo con los demás. Por ejemplo, no se puede definir un tipo de datos de intervalo que utilice el campo YEAR y el campo HOUR.

El tipo de datos de intervalo utiliza la palabra clave INTERVAL seguido por una cláusula de <calificador de intervalo>. La cláusula es una serie de reglas complejas que describen cómo se puede definir el tipo de datos de INTERVAL para expresar la participación de los intervalos de años, meses, días, horas, minutos o segundos. Además, el campo principal (la primera palabra) en la cláusula se puede definir con una precisión (*p*). La precisión es el número de dígitos que se utilizan en el campo principal. Si la precisión no se especifica, la precisión predeterminada es 2. Para los intervalos de año-mes, se puede especificar uno de los siguientes tipos de datos de intervalo:

- INTERVAL YEAR
- INTERVAL YEAR(*p*)
- INTERVAL MONTH
- INTERVAL MONTH(*p*)
- INTERVAL YEAR TO MONTH
- INTERVAL YEAR(*p*) TO MONTH

Hay muchas más opciones para los intervalos día-hora, ya que hay más campos de donde escoger. Por ejemplo, se pueden especificar cualesquiera de los siguientes tipos de intervalos utilizando el campo DAY como campo principal o independientemente del campo:

- INTERVAL DAY
- INTERVAL DAY(*p*)
- INTERVAL DAY TO HOUR
- INTERVAL DAY(*p*) TO HOUR
- INTERVAL DAY TO MINUTE
- INTERVAL DAY(*p*) TO MINUTE
- INTERVAL DAY TO SECOND
- INTERVAL DAY(*p*) TO SECOND
- INTERVAL DAY TO SECOND(*x*)
- INTERVAL DAY(*p*) TO SECOND(*x*)

Cuando el campo de rastreo (la última palabra) es `SECOND`, se puede especificar una precisión adicional (x), que define el número de dígitos después del punto decimal. Como se puede observar de estos ejemplos, hay muchos más tipos de datos de intervalos de día-hora que se pueden definir. Sin embargo, tenga en cuenta que el campo principal debe tener siempre una unidad de tiempo superior que el campo de rastreo. Por ejemplo, el campo `YEAR` es mayor que `MONTH`, y `hour` es mayor que `minute`.

Si fuera a utilizarse un tipo de datos de intervalo en una definición de columna, se vería algo como lo siguiente:

```
RANGO_FECHA INTERVAL YEAR(4) TO MONTH
```

En este ejemplo, un valor en esa columna incluiría cuatro dígitos para el año, un guión, y luego dos dígitos para el mes, tal como 1999-08. Si la precisión no se especifica para el año, el rango de año puede incluir sólo dos dígitos (00 hasta 99).

Tipo de datos booleanos

El tipo de datos booleano (a diferencia de los tipos de datos de intervalo) es muy sencillo y fácil de aplicar. El tipo de dato respalda una construcción de verdadero/falso que permite sólo tres valores: verdadero, falso o desconocido. Un valor nulo se evalúa como desconocido. (En SQL, un valor *nulo* se utiliza para indicar que un valor no está definido o no se conoce. Se analizarán los valores nulos en el capítulo 4.)

Los valores en el tipo de datos booleano se pueden utilizar en consultas de SQL y expresiones con fines de comparación. (Se analizarán las comparaciones en el capítulo 9.) Las comparaciones en el tipo de datos booleano siguen una lógica específica:

- Verdadero es mayor que falso.
- Una comparación con un valor desconocido (nulo) devolverá un resultado desconocido.
- Un valor desconocido se puede asignar a una columna sólo si admite valores nulos.

Para utilizar el tipo de datos booleano, debe utilizar la palabra clave `BOOLEAN` sin parámetros, como se muestra en el siguiente ejemplo:

```
ARTISTAS_CON_AGENTE BOOLEAN
```

La columna `ARTISTAS_CON_AGENTE` sólo aceptará valores de verdadero, falso y desconocido.

NOTA

El tipo de datos booleano se basa en un tipo específico de lógica del ordenador conocido como booleana (nombrado por el matemático del siglo XIX George Boole), que evalúa las condiciones de verdadero o falso en una operación o expresión determinada. Muchos lenguajes de programación respaldan la lógica booleana mediante el uso de operadores lógicos como `AND`, `OR` y `NOT`, por ejemplo, "`ARTÍCULO_A IS NOT FALSE`" o "`ARTÍCULO_A AND ARTÍCULO_B OR ARTÍCULO_C IS TRUE`". En SQL, la lógica booleana se aplica mediante el uso de operadores de comparación para equiparar valores dentro de varios tipos de datos. Estos operadores se analizarán en el capítulo 9.

Pregunta al experto

P: ¿Cómo se pueden comparar los tipos de datos predefinidos en SQL con los tipos de datos que se encuentran en otros lenguajes de programación?

R: En su mayor parte, es poco probable que los tipos de datos de dos diferentes lenguajes sean los mismos. Un conjunto de tipos de datos en un lenguaje pueden variar en estructura y semántica de un conjunto de tipos de datos en otro lenguaje. Estas diferencias, a menudo llamadas desajuste de impedancia, pueden conducir a la pérdida de información cuando una aplicación toma datos de una base de datos de SQL. De hecho, a menudo es una buena idea saber qué lenguaje se usará para las aplicaciones cuando se diseña la base de datos. En algunos casos, el diseño de la base de datos puede afectar el lenguaje de aplicación que más fácilmente se utiliza para manipular los datos en una base de datos de SQL. Sin embargo, SQL incluye una expresión de conversión de datos llamada CAST. La expresión CAST permite convertir datos de un tipo de datos a otro tipo de datos, permitiendo al lenguaje anfitrión acceder valores que no podrían manejarse en su forma original. La expresión CAST se analizará a mayor detalle en el capítulo 10.

P: ¿Se pueden asignar los tipos de datos de SQL a objetos que no sean columnas?

R: Cada valor de datos de SQL, o literal, pertenece a un tipo de datos. Por ejemplo, los tipos de datos se pueden asignar a los parámetros de los procedimientos invocados externamente. Los *procedimientos invocados externamente* son procedimientos que están contenidos dentro de un módulo de clientes de SQL. Un procedimiento es una instrucción SQL (o series de instrucciones) que pueden convocarse desde otro elemento en el código, que en el caso de procedimientos invocados externamente es un código externo. Un parámetro, que es el literal que pertenece a un tipo de datos, es un valor que se pasa al procedimiento y se utiliza cuando el procedimiento se procesa. El parámetro actúa como un marcador de posición para ese valor. Los módulos de clientes SQL se analizan en el capítulo 17.

Utilice tipos de datos SQL

Ahora que ha echado un vistazo a los diferentes tipos de datos predefinidos, veamos la instrucción CREATE TABLE, que define a una tabla con columnas que usan diferentes tipos de datos. En el siguiente ejemplo, la instrucción crea una tabla llamada ARTISTAS que incluye cuatro columnas:

```
CREATE TABLE ARTISTAS
( ID_ARTISTA          INT,
  NOMBRE_ARTISTA      VARCHAR(60) ,
  FDN_ARTISTA         DATE,
  POSTER_EN_EXISTENCIA  BOOLEAN );
```

ID_ARTISTA: INT	NOMBRE_ARTISTA: VARCHAR(60)	FDN_ARTISTA: DATE	POSTER_EN_EXISTENCIA: BOOLEAN
10001	Jennifer Warnes	1947-03-03	Falso
10002	Joni Mitchell	1943-11-07	Desconocido
10005	Bing Crosby	1904-05-02	Verdadero
10006	Patsy Cline	1932-09-08	Verdadero
10008	Placido Domingo	1941-01-21	Falso
10009	Luciano Pavarotti	1935-10-12	Desconocido

Figura 3-2 La tabla ARTISTAS definida con diferentes tipos de datos.

Como puede ver, la columna ID_ARTISTA tiene un tipo de datos numérico, la columna NOMBRE_ARTISTA tiene un tipo de datos de cadena, la columna FDN_ARTISTA tiene un tipo de datos de fecha y hora, y la columna POSTER_EN_EXISTENCIA tiene un tipo de datos booleano. La figura 3-2 muestra cómo puede verse esta tabla.

Creación de tipos definidos por el usuario

En el capítulo 1 se mencionó que el estándar SQL:2006 incorporó algunos de los principios de programación orientados a objetos (OOP) a su lenguaje. Un ejemplo de esto es el tipo definido por el usuario, a veces denominado como el tipo de datos definido por el usuario. El *tipo definido por el usuario* es una forma de tipo de datos (almacenada como un objeto de esquema) que está en parte definida por el programador y en parte basada en uno o más tipos de datos. SQL apoya dos formas de tipos definidos por el usuario:

- **Tipos estructurados** Estos tipos se forman por uno o más atributos, cada uno de los cuales está basado en otro tipo de datos, incluyendo tipos predefinidos, tipos construidos, y otros tipos estructurados. Además de asociarse con un tipo de datos, cada atributo puede incluir una cláusula predeterminada y puede especificar un cotejo. Un tipo estructurado puede incluir métodos en su definición. Un *método* es un tipo de función que se asocia con un tipo definido por el usuario. Una *función* es el nombre de una operación que realiza tareas predefinidas que no se podrían realizar normalmente mediante el uso de sólo instrucciones SQL. Es un tipo de ruta que toma parámetros de entrada (que a menudo son opcionales) y devuelve un valor único basado en esos parámetros.
- **Tipos distintos** Estos tipos se basan simplemente en tipos de datos predefinidos, y cualquier tipo de parámetro se define para ese tipo de datos si los parámetros son necesarios o deseados.

SQL proporciona la instrucción `CREATE TYPE` para definir los tipos definidos por el usuario. Sin embargo, el lenguaje utilizado para crear un tipo definido por el usuario puede variar de un producto a otro. Las características que se respaldan en el tipo definido por el usuario también varían ampliamente. Por ejemplo, SQL Server 2000 no respalda la instrucción `CREATE TYPE`, pero SQL Server 2005 sí lo hace.

A pesar de las diferencias y limitaciones en las aplicaciones de los productos, se trata por lo menos de proporcionar un ejemplo de cómo la instrucción `CREATE TYPE` se utiliza para crear un tipo diferente. En la siguiente instrucción se creó un tipo definido por el usuario que se basa en el tipo de datos `NUMERIC`:

```
CREATE TYPE SALARIO AS NUMERIC(8,2)
FINAL;
```

Este sencillo ejemplo es bastante claro: la creación de un tipo llamado `SALARIO` con un tipo de datos `NUMERIC(8,2)`. Sin embargo, la palabra clave `FINAL` es probablemente nueva. Cuando se especifica `FINAL`, le dice a SQL que no se definen subtipos para este tipo. La alternativa es especificar `NOT FINAL`, lo que significa que se pueden definir los subtipos para este tipo. Una vez creado el tipo, se puede utilizar en la definición de una columna como si fuera un tipo de datos predefinido:

```
CREATE TABLE EMPLEADO
( ID_EMPLEADO          INTEGER,
  SALARIO_EMPLEADO     SALARIO );
```

Cualquier valor que se agregue a la columna `SALARIO_EMPLEADO` se ajustará a las especificaciones del tipo de datos `NUMERIC` con una precisión de 8 y una escala de 2. Como resultado, el valor puede ser cualquiera desde -999999.99 hasta 999999.99. La mejor parte es que se puede utilizar el tipo definido por el usuario `SALARIO` en cualquier otro tipo de tablas que requieren valores similares.

NOTA

Como se observa, los tipos numéricos permiten números negativos, así como el número cero y números positivos. Si no se desean números negativos (que sería evidentemente el caso del salario de alguien), el tipo de datos por sí solo no haría el trabajo, pero se puede utilizar la restricción `CHECK` para ese fin. Las restricciones `CHECK` se analizarán en el capítulo 4.

Especificación de los valores predeterminados en una columna

Otra característica valiosa que SQL respalda es la habilidad de especificar un valor predeterminado para una columna cuando se utilice la instrucción `CREATE TABLE` para crear una tabla. La sintaxis para la definición de una columna con un valor predeterminado tendría este aspecto:

```
<nombre de columna> <tipo de datos> DEFAULT <valor predeterminado>
```

Los marcadores de posición <nombre de columna> y <tipo de datos>, con los cuales ahora debe estar familiarizado, son seguidos por la palabra clave DEFAULT. Después de la palabra clave DEFAULT, debe especificarse un valor para el marcador de posición <valor predeterminado>. Este valor puede ser un literal, que es un valor de datos de SQL (tal como la cadena 'A determinar' o el número 0); una función de valor de fecha y hora, que es una función que permite realizar operaciones relacionadas con fechas y tiempo (discutidas en el capítulo 10), o una función de usuario relacionada con una sesión, que es una función que devuelve la información relacionada con el usuario (discutida en el capítulo 10).

Cualquier tipo de valor utilizado para el marcador de posición <valor predeterminado> debe ajustarse a los requisitos de información del tipo de datos especificados en la definición de una columna. Por ejemplo, si se define una columna con un tipo de datos INT o un tipo de datos CHAR(4), no se puede especificar un valor predeterminado 'Desconocido'. En el primer caso, INT requiere un valor numérico, y en el segundo caso, CHAR(4) requiere que el valor contenga no más de cuatro caracteres.

En el siguiente ejemplo se utiliza la instrucción CREATE TABLE para definir a una tabla llamada ARTISTAS, que contiene tres columnas:

```
CREATE TABLE ARTISTAS
( ID_ARTISTA          INT,
  NOMBRE_ARTISTA      VARCHAR(60),
  LUGAR_DE_NACIMIENTO VARCHAR(60) DEFAULT 'Desconocido' );
```

Observe que la columna LUGAR_DE_NACIMIENTO incluye el valor predeterminado 'Desconocido'. El valor es aceptable, ya que se ajusta a los requisitos del tipo de datos VARCHAR(60). También observe que el valor predeterminado se encierra en comillas simples. Debe utilizar comillas simples para valores de cadena de caracteres. La figura 3-3 muestra cómo puede verse la tabla si se llena con filas de datos.

Si se insertan nuevas filas en esta tabla y no se sabe el lugar de nacimiento de un artista, el sistema automáticamente insertará el valor 'Desconocido'.

ID_ARTISTA: INT	NOMBRE_ARTISTA: VARCHAR(60)	LUGAR_DE_NACIMIENTO: VARCHAR(60)
10001	Jennifer Warnes	Desconocido
10002	Joni Mitchell	Fort MacLeod, Alberta, Canada
10005	Bing Crosby	Tacoma, Washington, Estados Unidos
10006	Patsy Cline	Winchester, Virginia, Estados Unidos
10008	Placido Domingo	Madrid, España
10009	Luciano Pavarotti	Desconocido

Figura 3-3 Un valor predeterminado 'Desconocido' para la columna LUGAR_DE_NACIMIENTO.

Pruebe esto 3-1 Creación de tablas en SQL

Probablemente notará que se ha utilizado información relacionada con CD para los ejemplos que se han mostrado hasta ahora. Se tratará este tema a lo largo del libro cuando se comience a construir la base de datos que registre el inventario de CD de una empresa pequeña. En este ejercicio se crean tres tablas que están relacionadas con la base de datos INVENTARIO, que se creó en el capítulo 2, Pruebe esto 2-1. Antes de empezar, eche un vistazo al modelo de datos simple (figura 3-4) que muestra las tres tablas que se crean. Cada tabla se representa por un rectángulo, con el nombre de la tabla en la parte superior del rectángulo y el nombre de las columnas, junto con los tipos de datos, que figuran dentro del rectángulo.

Se utilizará el modelo de datos a lo largo del libro (mientras evoluciona para convertirse en una estructura más compleja) para definir los objetos en la base de datos. También puede descargar el archivo Try_This_03.txt, que contiene instrucciones SQL utilizadas en este ejercicio (en inglés).

Paso a paso

1. Abra la aplicación de clientes de su RDBMS y conecte la base de datos INVENTARIO. Todos los objetos se crean dentro de esa base de datos. (Si el RDBMS no respalda la creación de una base de datos y en su lugar crea el esquema INVENTARIO_CD, debe crear todos los objetos dentro de ese esquema.)
2. La primera tabla que se crea es la tabla DISCOS_COMPACTOS. Observe que incluye tres columnas, dos de las cuales tienen un tipo de datos INT y una que tiene un tipo de datos VARCHAR(60). Esta tabla tiene información acerca de los discos compactos en el inventario. La columna ID_DISCO_COMPACTO contiene números que identifican únicamente a cada CD. La columna TITULO_CD contiene los nombres actuales de los CD. La columna ID_DISQUERA contiene números que identifican a las compañías que editan los CD. Introduzca la siguiente instrucción SQL en la ventana de entrada de la aplicación clientes:

```
CREATE TABLE DISCOS_COMPACTOS
( ID_DISCO_COMPACTO INT,
  TITULO_CD          VARCHAR(60) ,
  ID_DISQUERA        INT );
```

3. Verifique que se introdujo la información correcta y ejecute la instrucción. Debe recibir un mensaje de confirmación de que la instrucción se ejecutó exitosamente.

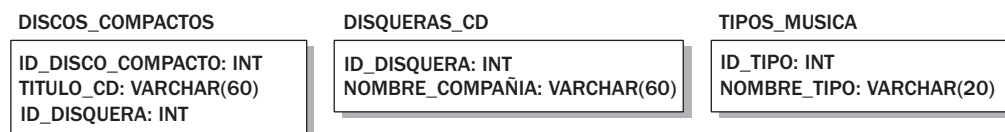


Figura 3-4 Modelo de datos simple de la base de datos INVENTARIO.

4. La próxima tabla que se crea es la tabla `DISQUERAS_CD`. La tabla incluye la columna `ID_DISQUERA`, que identifica únicamente a cada compañía que edita los CD, y la columna `NOMBRE_COMPañIA`, que enumera los nombres actuales de las compañías. Introduzca y ejecute el siguiente código:

```
CREATE TABLE DISQUERAS_CD
( ID_DISQUERA          INT,
  NOMBRE_COMPañIA      VARCHAR(60) );
```

5. La última tabla que se crea es la tabla `TIPOS_MUSICA`. La tabla incluye la columna `ID_TIPO`, que identifica únicamente cada categoría de música, y la columna `NOMBRE_TIPO`, que enumera los nombres actuales de las categorías de música (por ejemplo, blues o jazz). Introduzca y ejecute el siguiente código:

```
CREATE TABLE TIPOS_MUSICA
( ID_TIPO              INT,
  NOMBRE_TIPO          VARCHAR(20) );
```

6. Cierre la aplicación clientes.

Resumen de Pruebe esto

La base de datos ahora contiene tres tablas nuevas. Estas tablas sirven como base para otros ejercicios del libro. A medida que progrese mediante estos ejercicios, podrá modificar esas tablas, crear tablas adicionales, insertar datos en las tablas, y después hacer consultas y manipular esos datos. Para cuando complete todos los ejercicios, creará y completará una pequeña base de datos que almacene datos sobre un inventario de discos compactos.

Modificación de tablas en SQL

Tomando lo que aprendió acerca de la creación de tablas, puede utilizar la instrucción `ALTER TABLE` para modificar las definiciones de tablas base almacenadas en la base de datos (como objetos de esquema). En su forma más básica, la sintaxis para la instrucción `ALTER TABLE` se representa:

```
ALTER TABLE <nombre de la tabla>
  ADD [COLUMN] <definición de columna>
| ALTER [COLUMN] <nombre de columna>
  { SET DEFAULT <valor predeterminado> | DROP DEFAULT }
| DROP [COLUMN] <nombre columna> { CASCADE | RESTRICT }
```

La instrucción permite tomar tres diferentes acciones: añadir columnas, modificar columnas o eliminar columnas.

NOTA

La instrucción `ALTER TABLE` también permite añadir o eliminar limitaciones en una tabla. Una limitación en una tabla es una regla que restringe qué datos se introducen en la tabla. La limitación forma parte de la definición de la tabla, pero no forma parte de cualquiera de las definiciones de una columna específica. Las limitaciones se discuten con detalle en el capítulo 4.

El marcador de posición <definición de columna> en la cláusula `ADD [COLUMN]` es similar a la sección de definición de columna de la instrucción `CREATE TABLE`. Se proporciona un nombre de columna y un tipo de datos o dominio. También se tiene la opción de añadir una cláusula predeterminada, una limitación de columna o un cotejo. Por ejemplo, se puede utilizar la siguiente instrucción para modificar la tabla `ARTISTAS` de modo que incluya la columna `FDN_ARTISTA`:

```
ALTER TABLE ARTISTAS
ADD COLUMN FDN_ARTISTA DATE;
```

A diferencia de la cláusula `ADD [COLUMN]`, la cláusula `ALTER [COLUMN]` se limita a dos acciones: establecer una predeterminación o eliminar una predeterminación (aunque existen aplicaciones de productos que permiten cambiar otras propiedades, tal como el tipo de datos o la precisión y la escala). Por ejemplo, suponga que la tabla `ARTISTS` incluye la columna `LUGAR_DE_NACIMIENTO`, pero no se define ninguna predeterminación para esa columna. Puede agregar una predeterminación utilizando la siguiente instrucción:

```
ALTER TABLE ARTISTAS
ALTER COLUMN LUGAR_DE_NACIMIENTO SET DEFAULT 'Desconocido';
```

También puede eliminar la predeterminación utilizando la siguiente instrucción:

```
ALTER TABLE ARTISTAS
ALTER COLUMN LUGAR_DE_NACIMIENTO DROP DEFAULT;
```

La cláusula final en la sintaxis (`DROP[COLUMN]`) proporciona dos opciones para eliminar una columna y los datos de la tabla: las palabras clave `CASCADE` y `RESTRICT`. Es posible que recuerde estas palabras clave de la discusión sobre la instrucción `DROP SCHEMA` en el capítulo 2. Si se especifica la opción `CASCADE`, la columna y los datos dentro de la columna se eliminan independientemente de si otros objetos hacen referencia a esa columna. Todas las vistas, restricciones, rutas o activadores que hacen referencia a la columna también se eliminan. Si se utiliza la opción `RESTRICT`, la columna se elimina sólo si no hay vistas, restricciones, rutas o activadores que hagan referencia a la columna. Por ejemplo, la siguiente instrucción elimina la columna `LUGAR_DE_NACIMIENTO` y los datos almacenados en la columna, independientemente de las dependencias:

```
ALTER TABLE ARTISTAS
DROP COLUMN LUGAR_DE_NACIMIENTO CASCADE;
```

En general, es práctico conocer la instrucción `ALTER TABLE`, ya que las definiciones de las tablas cambian invariablemente, así como los tipos de datos almacenados en esas tablas. Sin embargo, esta instrucción, como la mayoría de las instrucciones de SQL, puede variar ampliamente de una aplicación a otra en términos de cómo los detalles de una instrucción se aplican. Por ejemplo, SQL Server no respalda las palabras clave `CASCADE` y `RESTRICT`. En Oracle, `CASCADE` debe escribirse como `CASCADE CONSTRAINTS`, y `RESTRICT` (que es el comportamiento predeterminado) no se respalda explícitamente. Como siempre, asegúrese de comprobar la documentación del producto.

Eliminación de tablas en SQL

Como puede imaginar, el proceso de eliminación de una tabla y sus datos almacenados es muy sencillo. La siguiente sintaxis muestra lo fácil que resulta este proceso:

```
DROP TABLE <nombre de la tabla>{ CASCADE | RESTRICT }
```

La única decisión que necesita tomar cuando se elimina una tabla es si debe escoger la opción CASCADE o RESTRICT. Como en los anteriores ejemplos de sintaxis, las dos opciones determinan si debe eliminar la tabla y sus datos aun cuando la tabla hace referencia a otros objetos. Si se utiliza CASCADE, la tabla y sus datos se eliminan, junto con todas las vistas, restricciones, rutas o activadores que hacen referencia a la tabla. Si se utiliza RESTRICT, la tabla se elimina sólo si no existen dichas dependencias. (Como con la cláusula DROP COLUMN, SQL Server no respalda CASCADE o RESTRICT, y Oracle permite sólo CASCADE CONSTRAINTS.) Por ejemplo, la siguiente instrucción elimina la tabla ARTISTAS y los datos almacenados en la columna, independientemente de las dependencias:

```
DROP TABLE ARTISTAS CASCADE;
```

Pregunta al experto

P: ¿Cómo se pueden eliminar los datos en una tabla, pero no la propia definición de la tabla?

R: En lugar de usar la instrucción DROP TABLE, tiene que utilizar la instrucción DELETE. La instrucción DELETE elimina todas las filas de una tabla o elimina sólo filas específicas, tal como se defina en la instrucción. Esto no es lo mismo que la instrucción DROP TABLE, que elimina la definición de la tabla y los datos. Se analizará la instrucción DELETE con más detalle en el capítulo 8. Muchas aplicaciones de productos también proporcionan la instrucción TRUNCATE, que suministra una forma rápida y eficaz para borrar todos los datos de una tabla. Sin embargo, la instrucción TRUNCATE no se incluye en el estándar SQL:2006.

P: Se expuso que cuando un valor predeterminado se define por una columna, el valor se inserta automáticamente en la columna cuando se agrega una fila a la tabla, pero no se especifica un valor para esa columna en particular. ¿Qué pasa si la definición de columna no incluye una predeterminación y se trata de insertar esa fila?

R: La acción tomada depende de si se permiten valores nulos dentro de la columna. Un valor nulo significa que el valor no se conoce. Esto no es lo mismo que cero, en blanco o predeterminado. Si un valor nulo se presenta, entonces los datos no están disponibles. Por predeterminación, todas las columnas permiten valores nulos, aunque se puede ignorar dicha predeterminación (discutido en el capítulo 4). Si trata de insertar una fila sin definir un valor específico, un valor nulo se inserta en esa columna si la columna permite valores nulos. Si la columna no permite valores nulos, no podrá insertar una fila sin definir un valor específico para esa columna.

(continúa)

P: A menudo se escucha el término “índices” en relación con la creación de tablas SQL. ¿Cómo se crean los índices?

R: Curiosamente, el estándar SQL:2006 no respalda la creación y mantenimiento de índices y tampoco proporciona una definición o los menciona de ninguna otra forma. Para los que no estén familiarizados con ellos, un índice es un conjunto de valores y referencias de búsqueda (en una tabla auxiliar) que corresponden a las filas de una tabla. Los índices aceleran las consultas y mejoran el rendimiento, haciendo el acceso a los datos mucho más eficiente, al igual que utilizar el índice de un libro ayuda a encontrar las cosas más rápidamente que la búsqueda secuencial de páginas. Como resultado, casi todos los RDBMS respaldan alguna forma de indexación y, de hecho, son una parte importante de ese producto. Sin embargo, el método utilizado para aplicar la indexación es muy variable; por lo tanto, cada producto proporciona su propio sistema para crear y mantener los índices. Por ejemplo, la instrucción CREATE INDEX está disponible en la mayoría de los productos; sin embargo, la sintaxis para la instrucción puede variar considerablemente. Como siempre, asegúrese de revisar la documentación del producto.

Pruebe esto 3-2 Modifique y elimine tablas en SQL

A lo largo del ciclo de vida de casi cualquier base de datos, la probabilidad de que los requerimientos del negocio cambien y la modificación de las bases de datos es casi una conclusión inevitable. Como resultado, no cabe duda que se pasará por situaciones en que las definiciones de tablas se modifican o eliminan. En este ejercicio se creará una tabla, se eliminará, se volverá a crear, y luego cambiará por la eliminación de una columna. En el momento en que termine, se añadirá una tabla más a la base de datos INVENTARIO y utilizará esa tabla en el último ejercicio. Puede descargar el archivo Try_This_03.txt, que contiene las instrucciones SQL utilizadas en este ejercicio (en inglés).

Paso a paso

1. Abra la aplicación de cliente del RDBMS y conéctese a la base de datos INVENTARIO (o el esquema INVENTARIO_CD).
2. Cree una tabla llamada TIPOS_DISCO_COMPACTO. La tabla incluye la columna ID_DISCO_COMPACTO y la columna ID_TIPO. Ambas columnas se asignarán a un tipo de datos INT. Introduzca y ejecute el siguiente código:

```
CREATE TABLE TIPOS_DISCO_COMPACTO
( ID_DISCO_COMPACTO INT,
  ID_TIPO            INT );
```

3. Ahora elimine la tabla de la base de datos. Introduzca y ejecute el siguiente código:

```
DROP TABLE TIPOS_DISCO_COMPACTO CASCADE;
```

4. Ahora se volverá a crear la tabla que se hizo en el paso 2, sólo que esta vez se incluirá una tercera columna llamada TITULO_CD con un tipo de datos VARCHAR(60). Introduzca y ejecute el siguiente código:

```
CREATE TABLE TIPOS_DISCO_COMPACTO
( ID_DISCO_COMPACTO INT,
  TITULO_CD          VARCHAR(60) ,
  ID_TIPO            INT );
```

5. El siguiente paso es eliminar la columna TITULO_CD. Introduzca y ejecute el siguiente código:

```
ALTER TABLE TIPOS_DISCO_COMPACTO
  DROP COLUMN TITULO_CD CASCADE ;
```

6. La tabla TIPOS_DISCO_COMPACTO debe ahora contener sólo las columnas ID_DISCO_COMPACTO y la columna ID_TIPO. Cierre la aplicación clientes.

Resumen de Pruebe esto

La base de datos INVENTARIO ahora debe contener cuatro tablas: DISCOS_COMPACTOS, DISQUERAS_CD, TIPOS_MUSICA y TIPOS_DISCO_COMPACTO. La tabla TIPOS_DISCO_COMPACTO, que se acaba de crear, contiene dos columnas, ID_DISCO_COMPACTO y ID_TIPO, ambas definidas con un tipo de datos INT. Estos ejercicios continuarán la construcción de esta base de datos agregando tablas nuevas y modificando las existentes.

Autoexamen Capítulo 3

1. ¿Qué tipos de tablas base se pueden crear utilizando una instrucción CREATE TABLE?
 - A Tablas base persistentes
 - B Tablas base temporales globales
 - C Tablas temporales locales creadas
 - D Tablas temporales locales declaradas
2. ¿Cuál es la principal diferencia entre una tabla temporal global y una tabla temporal local creada?
3. Está creando una tabla llamada AGENTES. La tabla incluye la columna ID_AGENTE, que tiene un tipo de datos INT, y la columna NOMBRE_AGENTE, que tiene un tipo de datos CHAR(60). ¿Qué instrucción SQL utilizaría?
4. ¿Cuáles son los tres tipos de datos que soporta SQL?
5. ¿Cuáles son los cuatro tipos de datos de cadena?

6. Un(a) _____ es un tipo de datos que permite valores que se basan en bits de datos, en lugar de conjuntos de caracteres o cotejos. Este tipo de datos permite sólo valores de 0 y 1.
7. ¿Cuál es la precisión y la escala del número 5293.472?
8. ¿Cuáles son las diferencias entre los tipos de datos numéricos exactos y los tipos de datos numéricos aproximados?
9. ¿Cuáles tipos de datos son tipos de datos numéricos exactos?
 - A DOUBLE PRECISION
 - B DECIMAL
 - C REAL
 - D SMALLINT
10. Un tipo de datos _____ especifica los valores de una fecha por año, mes y día.
11. ¿Cuáles son las dos formas de tipos de datos de intervalo que soporta SQL?
12. ¿Qué tipo de datos debe utilizarse para soportar una construcción verdadero/falso que pueda ser utilizada para comparar valores?
13. Está creando un tipo definido por el usuario distinto llamado CIUDAD. El tipo de usuario se basa en el tipo de datos CHAR(40). ¿Qué instrucción SQL utilizaría?
14. Se crea una tabla llamada CLIENTES. La tabla incluye la columna NOMBRE_CLIENTE y la columna CIUDAD_CLIENTE. Ambas columnas tienen un tipo de datos VARCHAR(60). La columna CIUDAD_CLIENTE también tiene el valor predeterminado *Seattle*. ¿Qué instrucción SQL utilizaría?
15. ¿Qué instrucción SQL deberá utilizarse para eliminar una columna de una tabla existente?
16. ¿Qué instrucción SQL deberá utilizarse para eliminar la definición de una tabla y todos los datos de SQL de una base de datos?
17. Una base de datos incluye una tabla llamada CANTANTES_OPERA. Se quiere agregar una columna llamada NACIONALIDAD a esa tabla. La columna debe tener el tipo de datos VARCHAR(40). ¿Qué instrucción SQL utilizaría?
18. Se desea eliminar la definición de la tabla CANTANTES_OPERA de la base de datos. También se quieren eliminar todos los datos y cualquier dependencia de la tabla. ¿Qué instrucción SQL utilizaría?

Capítulo 4

Implementación de la integridad de datos

Habilidades y conceptos clave

- Entienda las restricciones de integridad
 - Utilice restricciones NOT NULL
 - Añada restricciones UNIQUE
 - Añada restricciones PRIMARY KEY
 - Añada restricciones FOREIGN KEY
 - Defina restricciones CHECK
-

Una base de datos SQL debe hacer más que sólo almacenar datos. Se debe asegurar que el almacenamiento de los datos es el correcto. Si la integridad de los datos se compromete, los datos pueden ser inexactos o inconsistentes, poniendo en cuestionamiento la fiabilidad de la base de datos. Con el fin de asegurar la integridad de los datos, SQL proporciona una serie de *restricciones de integridad*, reglas que se aplican a la base de datos para restringir los valores que se pueden colocar en esas tablas. Se pueden aplicar restricciones a columnas individuales, a tablas individuales o a múltiples tablas. En este capítulo se discute cada tipo de restricción y se explica cómo se pueden aplicar a la base de datos SQL.

Entienda las restricciones de integridad

Las restricciones de integridad de SQL, a las que se les conoce simplemente como restricciones, pueden dividirse en tres categorías:

- **Restricciones relacionadas con tablas** Un tipo de restricción que se precisa dentro de la definición de una tabla. Las restricciones definidas a nivel de tabla pueden aplicar a una o más columnas.
- **Afirmaciones** Un tipo de restricción que se precisa dentro de una definición de afirmación (separado de la definición de una tabla). Una afirmación puede relacionarse con una o más tablas.
- **Restricciones de dominio** Un tipo de restricción que se precisa dentro de una definición de dominio (separado de la definición de una tabla). Una restricción de dominio se asocia con cualquier columna que se define dentro del dominio específico.

De estas tres categorías de restricciones, las restricciones relacionadas con tablas son las más comunes e incluyen el mayor número de opciones de restricción. Las restricciones relacionadas con tablas se dividen en dos subcategorías: restricciones de tabla y restricciones de columna. Las restricciones en ambas subcategorías se precisan en la definición de la tabla. Una restricción de columna se incluye con la definición de la columna, y una restricción de tabla se incluye como un elemento de la tabla, similar a la manera en que las columnas se definen como elementos de la tabla. (En el capítulo 3 se analizaron los elementos de tablas y las definiciones de columnas.) Tanto

las restricciones de columna como las restricciones de tabla soportan una serie de diferentes tipos de restricciones. Éste no es el caso para las restricciones de afirmación y dominio, que se limitan sólo a un tipo de restricción. La figura 4-1 proporciona una descripción general de los tipos de restricciones que se pueden crear.

En la parte superior de la ilustración se pueden observar las tres categorías de restricciones. Debajo de la categoría de restricciones relacionadas con tablas están las subcategorías restricciones de columna y restricciones de tabla, cada una de las cuales contienen tipos específicos de restricciones. Por ejemplo, las restricciones de tabla pueden incluir restricciones únicas (restricciones UNIQUE y de PRIMARY KEY), referenciales (restricciones FOREIGN KEY) y CHECK, mientras que las restricciones de columna pueden incluir la restricción NOT NULL, así como restricciones únicas, referenciales y CHECK. Sin embargo, las de dominios y afirmaciones sólo soportan las restricciones CHECK.

NOTA

En algunos lugares, el estándar SQL:2006 utiliza el término “restricción de tabla” para referirse a ambos tipos de restricciones relacionadas con tablas. Utilizaré el término “relacionado con tablas” para evitar confusión.

Como se muestra en la figura 4-1 existen cinco diferentes tipos de restricciones: NOT NULL, UNIQUE, PRIMARY KEY, FOREIGN KEY y CHECK. En SQL, las restricciones UNIQUE y PRIMARY KEY se consideran restricciones únicas, y las restricciones de FOREIGN KEY se consideran como restricciones referenciales. El resto del capítulo se dedica a explicar lo que significan cada una de estas restricciones y cómo aplicarlas.

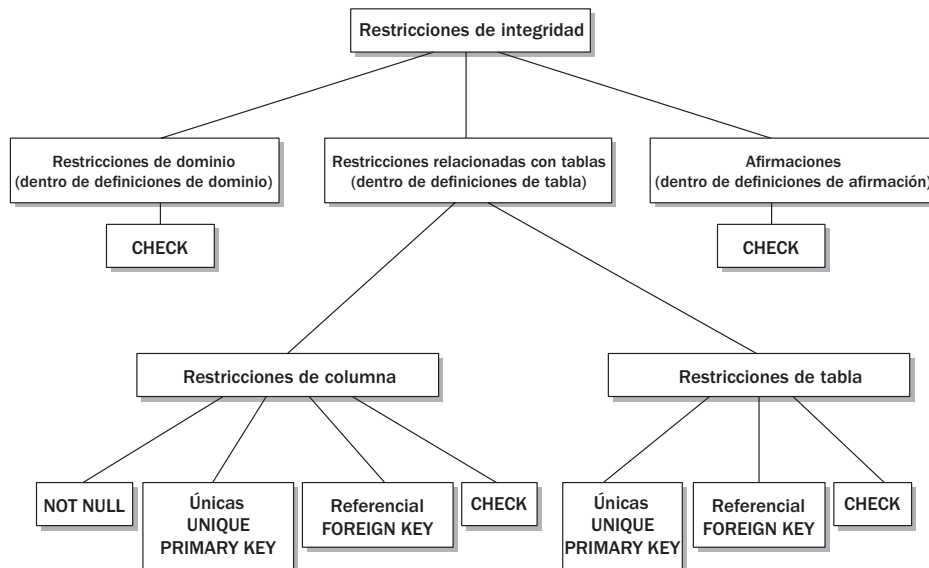


Figura 4-1 Tipos de restricciones de integridad de SQL.

Utilice restricciones NOT NULL

En el capítulo 3 se dijo que *nulo* significa que un valor no está definido o no se conoce. Esto no es lo mismo que cero, en blanco, una cadena vacía o un valor predeterminado. En lugar de ello, indica que un valor de dato está ausente. Se puede pensar en un valor nulo como una marca. (Una *marca* es un carácter, número o bit que indica un cierto hecho sobre una columna. La marca sirve como un marcador que designa una condición particular o la existencia de algo.) En el caso nulo, si una columna no proporciona ningún valor, se coloca la marca, indicando que el valor es desconocido, o nulo. Cada columna tiene una característica *de nulo* que indica si la columna acepta valores nulos. Por predeterminación, todas las columnas aceptan valores nulos. Sin embargo, se puede anular la característica de nulo predeterminada utilizando una restricción NOT NULL, que indica que la columna *no* acepta valores nulos.

NOTA

Algunos RDBMS permiten modificar la característica de nulo predeterminada de cualquier columna nueva que se crea. Además, algunos sistemas respaldan una restricción NULL, que se utiliza para designar que una columna acepta valores nulos.

La restricción NOT NULL sólo puede utilizarse como una restricción de columna. No se soporta para restricciones de tablas, afirmaciones o restricciones de dominio. La aplicación de una restricción NOT NULL es un proceso muy sencillo. Simplemente utilice la siguiente sintaxis cuando se crea una definición de columna:

```
<nombre de columna> { <tipo de datos> | <dominio> } NOT NULL
```

Por ejemplo, suponga que quiere crear una tabla llamada ARTISTAS_DISCO_COMPACTO que requiere tres columnas: ID_ARTISTA, NOMBRE_ARTISTA y LUGAR_DE_NACIMIENTO. Quiere asegurarse de que cualquier fila que se agregue a la tabla incluya un valor para la columna ID_ARTISTA y un valor para la columna NOMBRE_ARTISTA. Para hacer esto, debe añadir una restricción NOT NULL en ambas definiciones de columna, como se muestra en la siguiente instrucción de SQL:

```
CREATE TABLE ARTISTAS_DISCO_COMPACTO
( ID_ARTISTA          INT          NOT NULL,
  NOMBRE_ARTISTA      VARCHAR(60)  NOT NULL,
  LUGAR_DE_NACIMIENTO VARCHAR(60) );
```

Observe que la columna LUGAR_DE_NACIMIENTO no incluye una restricción NOT NULL. Como resultado, si no se suministra un valor para esa columna (cuando se inserta una fila), un valor nulo se insertará. (La marca se coloca.) La figura 4-2 muestra el aspecto que puede tener la tabla si se insertan filas que no contienen valores para la columna LUGAR_DE_NACIMIENTO.

Como puede verse, las columnas ID_ARTISTA, NOMBRE_ARTISTA no contienen (y no pueden contener) valores nulos. La columna LUGAR_DE_NACIMIENTO, por otro lado, contiene dos valores nulos.

ID_ARTISTA: INT:	NOMBRE_ARTISTA: VARCHAR(60)	LUGAR_DE_NACIMIENTO: VARCHAR(60)
10001	Jennifer Warnes	NULL
10002	Joni Mitchell	Fort MacLeod, Alberta, Canada
10005	Bing Crosby	Tacoma, Washington, Estados Unidos
10006	Patsy Cline	Winchester, Virginia, Estados Unidos
10008	Placido Domingo	Madrid, España
10009	Luciano Pavarotti	NULL

Figura 4-2 Valores nulos que aparecen en la columna LUGAR_DE_NACIMIENTO en la tabla ARTISTAS_DISCO_COMPACTO.

Añada restricciones UNIQUE

Si se remite de nuevo a la figura 4-1, notará que tanto las restricciones de columna como las restricciones de tabla respaldan restricciones únicas. También se observa que hay dos tipos de restricciones únicas: UNIQUE y de PRIMARY KEY. Esta sección se enfoca en la restricción UNIQUE. La restricción de PRIMARY KEY se analizará en la sección “Añada restricciones PRIMARY KEY”, más adelante en este capítulo.

La restricción UNIQUE permite exigir que una columna o conjunto de columnas contengan valores únicos, valores significativos que sean diferentes de todas las demás filas en la misma tabla. Por ejemplo, eche un vistazo a la figura 4-3, que muestra la tabla CD_INVENTARIO. La tabla contiene tres columnas: NOMBRE_ARTISTA, NOMBRE_CD y DERECHOSDEAUTOR.

NOMBRE_ARTISTA: VARCHAR(40)	NOMBRE_CD: VARCHAR(60)	AÑO_DERECHOSDEAUTOR: INT
Jennifer Warnes	Famous Blue Raincoat	1991
Joni Mitchell	Blue	1971
William Ackerman	Past Light	1983
Kitaro	Kojiki	1990
Bing Crosby	That Christmas Feeling	1993
Patsy Cline	Patsy Cline: 12 Greatest Hits	1988

Figura 4-3 Tabla CD_INVENTARIO con las columnas NOMBRE_ARTISTA, NOMBRE_CD y AÑO_DERECHOSDEAUTOR.

Puede decidir que desea que los valores en la columna `NOMBRE_CD` sean únicos, de modo que no haya dos nombres iguales de CD. Si se aplica una restricción `UNIQUE` a la columna, no será posible insertar una fila que contenga un valor de `NOMBRE_CD` que ya exista en la tabla. Suponga que ahora se da cuenta que hacer los valores `NOMBRE_CD` únicos no es una buena idea, ya que es posible para más de un CD compartir el mismo nombre. Decide adoptar otro enfoque y utilizar una restricción `UNIQUE` en las columnas `NOMBRE_ARTISTA` y `NOMBRE_CD`. De esa manera, ningún par `NOMBRE_ARTISTA/NOMBRE_CD` se puede repetir. Se puede repetir el valor `NOMBRE_ARTISTA` o el valor `NOMBRE_CD`, pero no se puede repetir la misma combinación de los dos. Por ejemplo, la tabla ya contiene la fila con el valor `NOMBRE_ARTISTA` de Joni Mitchell y el valor `NOMBRE_CD` de Blue. Si una restricción `UNIQUE` se aplica a estas dos columnas, no se podrá añadir otra fila que contenga *ambos* valores.

NOTA

Cabe señalar que las tablas utilizadas para ilustrar los conceptos en este capítulo no son necesariamente buenos diseños. Por ejemplo, los nombres de las personas y cosas rara vez son buenas opciones para identificar únicamente las filas de los datos porque son bastante largos (comparado con números), tienden a cambiar, y son propensos a problemas con valores duplicados. Sin embargo, estas tablas se escogieron ya que ilustran bien los conceptos.

Ahora que se tiene un entendimiento básico de cómo las restricciones `UNIQUE` se aplican, echemos un vistazo a la sintaxis que se utiliza para crearlas. Recuerde que se dijo que se puede crear una restricción `UNIQUE` que sea una restricción de columna o una restricción de tabla. Para crear una restricción de columna, añádala como parte de la definición de una columna, como se muestra en la siguiente sintaxis:

```
<nombre de columna> { <tipo de datos> | <dominio> } UNIQUE
```

Si se quiere añadir una restricción única como una restricción de tabla, se debe agregar como un elemento de definición de tabla, como se muestra en la siguiente sintaxis:

```
[ CONSTRAINT <nombre de la restricción> ]  
UNIQUE ( <nombre de columna> [{, <nombre de columna>}... ] )
```

Como se puede observar, aplicar una restricción `UNIQUE` como una restricción de columna es un poco más simple que aplicarla como una restricción de tabla. Sin embargo, si se aplica la restricción al nivel de columna, se puede aplicar sólo a una columna. A pesar de si se utiliza una restricción de columna o de tabla, se pueden definir tantas restricciones `UNIQUE` como sean necesarias en una sola definición de tabla.

Ahora volvamos a la tabla de la figura 4-3 que se usará para crear ejemplos de códigos para aplicar restricciones `UNIQUE`. En el primer ejemplo, se aplicó una restricción `UNIQUE` a la columna `NOMBRE_CD`:

```
CREATE TABLE CD_INVENTARIO  
( NOMBRE_ARTISTA  VARCHAR(40) ,  
  NOMBRE_CD       VARCHAR(60) UNIQUE ,  
  DERECHOSDEAUTOR INT  );
```

También se pueden aplicar restricciones UNIQUE a otras columnas, pero no tendrían el mismo efecto que la combinación de dos columnas en una restricción de tabla, como se muestra en el siguiente ejemplo:

```
CREATE TABLE CD_INVENTARIO
( NOMBRE_ARTISTA VARCHAR(40),
  NOMBRE_CD       VARCHAR(60),
  DERECHOSDEAUTOR INT,
  CONSTRAINT UN_ARTISTA_CD UNIQUE ( NOMBRE_ARTISTA, NOMBRE_CD ) );
```

Las columnas NOMBRE_ARTISTA y NOMBRE_CD ahora deben contener combinaciones únicas de valores con el fin de que una fila se añada a la tabla CD_INVENTARIO.

Hasta ahora se ha dicho que una restricción UNIQUE evita la duplicación de valores que se introducen en una columna o columnas definidas con esa restricción. Sin embargo, hay una excepción para esto (el valor nulo). Una restricción UNIQUE permite múltiples valores nulos en una columna. Como con otras columnas, los valores nulos se permiten por predeterminación. Sin embargo, se puede anular la predeterminación utilizando una restricción NOT NULL conjuntamente con una restricción UNIQUE. Por ejemplo, se puede añadir una restricción NOT NULL a la definición de columna NOMBRE_CD:

```
CREATE TABLE CD_INVENTARIO
( NOMBRE_ARTISTA VARCHAR(40),
  NOMBRE_CD       VARCHAR(60) NOT NULL UNIQUE,
  DERECHOSDEAUTOR INT );
```

También se puede añadir una restricción NOT NULL a la definición de una columna que hace referencia a una restricción de tabla:

```
CREATE TABLE CD_INVENTARIO
( NOMBRE_ARTISTA VARCHAR(40),
  NOMBRE_CD       VARCHAR(60) NOT NULL,
  DERECHOSDEAUTOR INT,
  CONSTRAINT UN_ARTISTA_CD UNIQUE (NOMBRE_CD) );
```

En cada caso, tanto la restricción NOT NULL como la restricción UNIQUE se aplican en la columna NOMBRE_CD, lo que significa que los valores NOMBRE_CD deben ser únicos y sin valores nulos.

Añada restricciones PRIMARY KEY

Como se mencionó en la sección “Añada restricciones UNIQUE”, una restricción PRIMARY KEY, como una restricción UNIQUE, es un tipo de restricción única de SQL. Ambos tipos de restricciones permiten sólo valores en columnas específicas, ambos tipos pueden aplicarse a una o más columnas, y ambos tipos se pueden definir como restricciones de columna o restricciones de tabla. Sin embargo, las restricciones PRIMARY KEY tienen dos restricciones que aplican sólo a éstas:

- Una columna que se define con una restricción PRIMARY KEY no puede contener valores nulos. No importa si la definición de columna especifica NOT NULL (la columna no puede contener valores nulos, ya que es una restricción PRIMARY KEY).
- Sólo una restricción PRIMARY KEY puede definirse para cada tabla.

El motivo de estas restricciones es el rol que una restricción de clave primaria (identificador único) juega en la tabla. Como podría recordar del capítulo 1, cada fila en una tabla debe ser única. Esto es importante, ya que SQL no puede diferenciar entre dos filas que son completamente idénticas; por lo tanto, no se puede actualizar o eliminar una fila duplicada sin hacer lo mismo con la otra. La clave primaria para una tabla se escoge por el diseñador de la base de datos de claves de candidato disponibles. Una *clave de candidato* es un conjunto de una o más columnas que identifican de forma exclusiva a cada fila. Por ejemplo, en la figura 4-4, la única clave de candidato razonable en la tabla CD_ARTISTAS es la columna ID_ARTISTA. Cada valor en la columna es único. De esa manera, aun si los valores NOMBRE_ARTISTA y AGENCIAS se duplican, la fila sigue siendo única ya que el valor ID_ARTISTA siempre es único.

La singularidad de una clave de candidato se puede ejecutar con una restricción UNIQUE o con una restricción PRIMARY KEY. Sin embargo, cada tabla debe incluir una clave primaria aun si no se define ninguna restricción UNIQUE. Esto se considera como una de las mejores prácticas en la industria, ya que una clave primaria no puede aceptar valores nulos, lo que la hace la medida definitiva por la cual la singularidad de una fila se puede asegurar. Las claves primarias también son útiles cuando una tabla hace referencia de otra a través del uso de claves foráneas. (Ver la sección “Añada restricciones FOREIGN KEY” más adelante en este capítulo.) Además, algunos RDBMS requieren la definición de claves primarias en determinadas circunstancias, como cuando una columna de tabla se incluye en un índice de texto completo.

Para definir una clave primaria, se debe utilizar una restricción de PRIMARY KEY para especificar qué columna o columnas servirán como la clave primaria de una tabla. El proceso de definición de una restricción de PRIMARY KEY es muy similar a la de la definición de la restricción



ID_ARTISTA: INT	NOMBRE_ARTISTA: VARCHAR(60)	ID_AGENCIA: INT
10001	Jennifer Warnes	2305
10002	Joni Mitchell	2306
10003	William Ackerman	2306
10004	Kitaro	2345
10005	Bing Crosby	2367
10006	Patsy Cline	2049
10007	Jose Carreras	2876
10008	Placido Domingo	2305
10009	Luciano Pavarotti	2345

Figura 4-4 La clave de candidato en la tabla CD_ARTISTAS.

UNIQUE. Si se quiere añadir una restricción PRIMARY KEY a una definición de columna, utilice la siguiente sintaxis:

<nombre de columna> { <tipo de datos> | <dominio> } PRIMARY KEY

Si se quiere añadir una restricción PRIMARY KEY como una restricción de tabla, se debe agregar como un elemento de tabla en la definición de la tabla, como se muestra en la siguiente sintaxis:

[CONSTRAINT <nombre de la restricción>]

PRIMARY KEY (<nombre de columna> [{, <nombre de columna> }...])

Como con la restricción UNIQUE, se puede utilizar una restricción de columna para definir una clave primaria si se incluye sólo una columna en la definición. Por ejemplo, si se define una restricción de PRIMARY KEY para la tabla mostrada en la figura 4-4, se utiliza la siguiente instrucción SQL:

```
CREATE TABLE CD_ARTISTAS
( ID_ARTISTA      INT              PRIMARY KEY,
  NOMBRE_ARTISTA  VARCHAR(60) ,
  ID_AGENCIA      INT ) ;
```

Si se quiere aplicar la restricción a varias columnas (o si simplemente se quiere mantener como una definición separada, entonces se debe utilizar la restricción de tabla:

```
CREATE TABLE CD_ARTISTAS
( ID_ARTISTA      INT,
  NOMBRE_ARTISTA  VARCHAR(60) ,
  ID_AGENCIA      INT,
  CONSTRAINT PK_ID_ARTISTA PRIMARY KEY ( ID_ARTISTA, NOMBRE_ARTISTA ) ) ;
```

Este método crea una clave primaria en las columnas ID_ARTISTA y NOMBRE_ARTISTA, de modo que los valores combinados de ambas columnas deben ser únicos, aunque puede existir duplicación dentro de una columna individual. Un experimentado diseñador de bases de datos señalará rápidamente que ésta es una *superclave*, lo que significa que tiene más columnas que el mínimo necesario para formar una clave primaria. Y esto es cierto (ID_ARTISTA por sí misma es única, y realmente no es necesario añadirle NOMBRE_ARTISTA con el fin de formar una clave primaria, y si se quiere estar seguro de que los valores duplicados de ID_ARTISTA no se incluyen en la tabla, lo que significa que se debe tener una clave primaria con sólo un ID_ARTISTA en ella). Se hace aquí para ilustrar que una clave primaria puede contener múltiples columnas y para definir una forma que una restricción de tabla debe utilizar.

Puede encontrar que desea definir tanto una restricción PRIMARY KEY como una restricción UNIQUE en una tabla. Para hacerlo, simplemente defina las restricciones como normalmente lo hace. Por ejemplo, la siguiente instrucción de SQL define una restricción PRIMARY KEY en la columna ID_ARTISTA y una restricción UNIQUE en la columna NOMBRE_ARTISTA:

```
CREATE TABLE CD_ARTISTAS
( ID_ARTISTA      INT              PRIMARY KEY,
  NOMBRE_ARTISTA  VARCHAR(60) ,
  ID_AGENCIA      INT,
  CONSTRAINT UN_NOMBRE_ARTISTA UNIQUE (NOMBRE_ARTISTA) ) ;
```

Se alcanza el mismo resultado con el siguiente código:

```
CREATE TABLE CD_ARTISTAS
( ID_ARTISTA      INT,
  NOMBRE_ARTISTA  VARCHAR(60) UNIQUE,
  ID_AGENCIA      INT,
  CONSTRAINT PK_ID_ARTISTA PRIMARY KEY (ID_ARTISTA) );
```

NOTA

Se utiliza una restricción UNIQUE en estas instrucciones de SQL sólo como una forma de demostrar cómo se puede usar la restricción en una tabla con una clave primaria. Lo más probable es que no se quiera utilizar una restricción UNIQUE para la columna NOMBRE_ARTISTA, ya que es posible para dos artistas compartir el mismo nombre. (Por ejemplo, dos diferentes artistas de blues, ambos de los cuales vivieron en la primera parte del siglo pasado, con el nombre de Sonny Boy Williamson.)

Pregunta al experto

- P:** ¿Pueden las columnas en una tabla pertenecer tanto a una restricción UNIQUE como a una restricción PRIMARY KEY?
- R:** Sí, siempre y cuando no sean exactamente las mismas columnas. Por ejemplo, suponga que tiene una tabla que incluye tres columnas: ID_ARTISTA, NOMBRE_ARTISTA y LUGAR_DE_NACIMIENTO. Se puede definir una restricción PRIMARY KEY que incluya las columnas ID_ARTISTA y NOMBRE_ARTISTA, que garantice pares de valores únicos en esas dos columnas, pero los valores dentro de columnas individuales pueden aún estar duplicados. Sin embargo, se puede entonces definir una restricción UNIQUE que incluya sólo la columna NOMBRE_ARTISTA para garantizar que esos valores también son únicos. (Ciertamente este no es el mejor diseño, pero ilustra el punto.) También se puede crear una restricción UNIQUE que incluya las columnas NOMBRE_ARTISTA y LUGAR_DE_NACIMIENTO para garantizar pares de valores únicos en esas dos columnas. Lo único que se puede hacer es crear una construcción UNIQUE que incluya exactamente las mismas columnas como en la restricción PRIMARY KEY y viceversa.
- P:** Se expuso que una columna que se incluye en una restricción PRIMARY KEY no acepta valores nulos. ¿Qué sucede si esa columna se configura también con una restricción NOT NULL?
- R:** Nada diferente sucede. La tabla se crea de la misma forma. Una definición de columna que incluye una PRIMARY KEY está diciendo lo mismo que una definición de columna que incluye NOT NULL PRIMARY KEY. De hecho, antes de SQL-92, las palabras clave NOT NULL se requerían en todas las columnas incluidas en una restricción PRIMARY KEY.

Cabe mencionar lo mismo para las restricciones UNIQUE. No fue sino hasta SQL-92 que se permitieron los valores nulos en columnas incluidas en una restricción UNIQUE, que las diferenciaban claramente de las restricciones PRIMARY KEY. También desconfíe de las diferentes implementaciones entre los proveedores. Por ejemplo, Oracle añade automáticamente restricciones NOT NULL a columnas incluidas en una restricción PRIMARY KEY, mientras que SQL Server (o por lo menos algunas versiones) expone un error si se intenta crear una restricción de PRIMARY KEY utilizando columnas que no tienen especificada una restricción NOT NULL.

Añada restricciones FOREIGN KEY

Hasta este punto, los tipos de restricciones que se analizaron tienen que ver principalmente con la garantía de la integridad de los datos dentro de una tabla. La restricción NOT NULL evita el uso de valores nulos dentro de una columna, y las restricciones UNIQUE y PRIMARY KEY garantizan la singularidad de los valores dentro de una columna o conjunto de columnas. Sin embargo, la restricción FOREIGN KEY es diferente en el sentido de que se ocupa de cómo los datos en una tabla hacen referencia a los datos en otra tabla, que es la razón por la que se conoce como una *restricción referencial* (en relación con otra tabla). (De hecho, hay una excepción, llamada *relación recursiva*, donde la clave foránea hace referencia a otra fila en la misma tabla, pero por el momento se ignorará este caso especial para centrarse en los fundamentos.)

Posiblemente recuerde del capítulo 1 que las tablas en una base de datos relacional están unidas entre sí de una manera significativa con el fin de garantizar la integridad de los datos. Esta asociación entre tablas forma una relación que proporciona una *integridad referencial* entre las tablas. La integridad referencial evita la manipulación de los datos en una tabla que afecte negativamente los datos en otra tabla. Echemos un vistazo a un ejemplo que ilustra este punto. La figura 4-5 muestra dos tablas (TITULOS_CD y EDITORES_CD) que están definidas con una clave primaria.

TITULOS_CD			EDITORES_CD	
ID_TITULO_CD: INT	TITULO_CD: VARCHAR(60)	ID_EDITOR: INT	ID_EDITOR: INT	NOMBRE_COMPAÑIA: VARCHAR(60)
11001	Famous Blue Raincoat	5422	5403	MCA Records
11002	Blue	5402	5402	Reprise Records
11003	Past Light	5412	5409	Geffen
11004	Kojiki	5409	5412	Windham Hill Records
11005	That Christmas Feeling	5403	5422	Private Music
11006	Patsy Cline: 12 Greatest Hits	5403		

Figura 4-5 Relación entre las tablas TITULOS_CD y EDITORES_CD.

La columna `ID_TITULO_CD` en la tabla `TITULOS_CD` se configuró con una restricción `PRIMARY KEY`, así como la columna `ID_EDITOR` en la tabla `EDITORES_CD`. Ambas columnas están sombreadas en la ilustración.

Observe que la tabla `TITULOS_CD` contiene una columna llamada `ID_EDITOR`. Esta columna incluye valores de la columna `ID_EDITOR` de la tabla `EDITORES_CD`. De hecho, los valores de `ID_EDITOR` en la tabla `TITULOS_CD` incluyen sólo valores que vienen de la columna `ID_EDITOR` en la tabla `EDITORES_CD`. No será posible insertar una fila en `TITULOS_CD` si el valor `ID_EDITOR` no figura en la tabla `EDITORES_CD`. Al mismo tiempo, si se modifica o elimina el valor `ID_EDITOR` en la tabla `EDITORES_CD`, se puede predecir el resultado de la acción si esos mismos valores existen en la tabla `TITULOS_CD`. Bajo ninguna circunstancia se querrá eliminar a un editor y dejar los valores `ID_EDITOR` en la tabla `TITULOS_CD` que hacen referencia a un editor que ya no existe. Estos resultados se pueden lograr utilizando una restricción de `FOREIGN KEY`. Una restricción `FOREIGN KEY` aplica la integridad referencial entre dos tablas para garantizar que no se tome ninguna acción en cualquiera de las tablas que afecte negativamente los datos protegidos por la restricción.

En las tablas mostradas en la figura 4-5, la restricción `FOREIGN KEY` debe configurarse en la columna `ID_EDITOR` de la tabla `TITULOS_CD`. La restricción `FOREIGN KEY` restringe los valores en esa columna de los valores de un clave de candidato (a menudo la clave primaria) en la tabla relacionada. Sólo se permiten los valores de datos válidos en la columna o columnas `FOREIGN KEY`.

NOTA

La tabla que contiene la clave foránea es la tabla de referencia. La tabla a la que se hace referencia en la clave foránea es la tabla referenciada. Asimismo, la columna o columnas que componen la clave foránea en la tabla referenciada hacen referencia a éstas como columnas referenciadas. Al hacer referencia a las columnas por la clave foránea, éstas son columnas referenciadas.

Cuando se crea una restricción `FOREIGN KEY`, se deben seguir varias directrices:

- Las columnas referenciadas se deben definir con una restricción `UNIQUE` o una `PRIMARY KEY`. Como podrá adivinar, la de clave primaria es la más utilizada por las columnas referenciadas.
- Una restricción `FOREIGN KEY` se puede crear como una restricción de tabla o una restricción de columna. Si se crea la clave foránea como una restricción de columna, se puede incluir sólo una columna. Si se crea la clave foránea como una restricción de tabla, se pueden incluir una o más columnas.
- La clave foránea en la tabla de referencia debe incluir el mismo número de columnas que sean referenciadas, y las columnas de referencia deben cada una configurarse con los mismos tipos de datos que su contraparte de referencia. Sin embargo, las columnas de referencia no deben tener necesariamente los mismos nombres que las columnas referenciadas.
- Si no se especifican las columnas referenciadas cuando se define una restricción `FOREIGN KEY`, entonces las columnas definidas en la clave primaria de la tabla referenciada se utilizan como columnas referenciadas.

Estas directrices estarán más claras cuando explique cómo implementar una clave foránea. En primer lugar, echemos un vistazo a la sintaxis básica utilizada para crear esa restricción. Si se quiere añadir una restricción FOREIGN KEY como una restricción de columna, se debe agregar la restricción a la definición de una columna, como se muestra en la siguiente sintaxis:

```
<nombre de columna> { <tipo de datos> | <dominio> } [NOT NULL]
REFERENCES <tabla referenciada> [( <columnas referenciadas> )]
[MATCH { FULL | PARTIAL | SIMPLE }]
[ < acción referencial desencadenada > ]
```

Si desea añadir una restricción FOREIGN KEY como restricción de tabla debe agregarla como elemento de tabla en la definición, como en la siguiente sintaxis:

```
[CONSTRAINT <nombre de la restricción>]
FOREIGN KEY (<columna referenciada> [{, <columna referenciada> }...])
REFERENCES <tabla referida> [(columnas de referencia>)]
[MATCH { FULL | PARTIAL | SIMPLE }]
[ < acción referencial desencadenada > ]
```

Como puede observarse, la restricción FOREIGN KEY es un poco más compleja que la sintaxis de la restricción que se había analizado hasta ahora. Sin embargo, crear la restricción básica FOREIGN KEY es un proceso relativamente simple. Echemos un vistazo a uno primero, para luego continuar con los elementos del lenguaje más complejos.

En el siguiente ejemplo se utiliza la instrucción CREATE TABLE para crear la tabla TITULOS_CD (mostrada en la figura 4-5) y definir una restricción de columna:

```
CREATE TABLE TITULOS_CD
( ID_TITULO_CD INT,
  TITULO_CD VARCHAR(60),
  ID_EDITOR INT REFERENCES EDITORES_CD );
```

Esta instrucción define la restricción FOREIGN KEY en la columna ID_EDITOR. Observe que, con el fin de añadir una restricción de columna, todo lo que se tiene que hacer es agregar la palabra clave REFERENCES y el nombre de la tabla referenciada. También note que la clave foránea contiene el mismo número de columnas que la clave primaria en la tabla referenciada, y las columnas referenciadas y de referencia son del mismo tipo de datos. Recuerde: si no se hace referencia de la clave primaria en la tabla referenciada, se debe también incluir el nombre de la columna o columnas (por ejemplo, REFERENCES EDITORES_CD (ID_EDITOR)).

NOTA

Antes de poder crear una clave foránea en una tabla, la tabla referenciada ya debe existir y se debe definir una restricción UNIQUE o PRIMARY KEY para esa tabla.

En el siguiente ejemplo se crea una clave foránea, que es una restricción de tabla. A diferencia del ejemplo anterior, se incluye el nombre de una columna referenciada en esta definición de restricción, aun cuando no es necesario:

```
CREATE TABLE TITULOS_CD
( ID_TITULO_CD INT,
  TITULO_CD VARCHAR(60),
  ID_EDITOR INT,
  CONSTRAINT FK_ID_EDITOR FOREIGN KEY (ID_EDITOR)
    REFERENCES EDITORES_CD (ID_EDITOR) );
```

Las últimas dos líneas del código son la definición de restricción. El nombre de la restricción, `FK_ID_EDITOR`, viene después de la palabra clave `CONSTRAINT`. Los nombres de las restricciones no son necesarios porque el RDBMS asigna un nombre generado por el sistema si no se proporciona uno. Sin embargo, es una buena práctica suministrar uno propio, ya que los nombres de las restricciones a menudo muestran mensajes de error cuando una instrucción SQL intenta violar una restricción, y los nombres suministrados serán fáciles de reconocer de aquellos proporcionados por el DBMS. Después del nombre de la restricción, las palabras clave `FOREIGN KEY` indican el tipo de restricción, que es seguido por el nombre de la columna de referencia, `ID_EDITOR`. Éste es el nombre de la columna en la que la restricción tendrá lugar. Si hubiera varios nombres de columnas, deben separarse por comas. El nombre de la columna de referencia es entonces seguido por la palabra clave `REFERENCES`, que es seguido por el nombre de la tabla referenciada, `EDITORES_CD`. El nombre de la columna referenciada viene después del nombre de la tabla referenciada.

Eso es todo lo que hay que hacer. Una vez que la restricción se define, no podrá poner valores en la columna `ID_EDITOR` de la tabla `TITULOS_CD` a menos que esos valores ya existan en la clave primaria de la tabla `EDITORES_CD`. Debe observar, sin embargo, que los valores en la clave foránea no tienen que ser únicos, como deben ser en la clave primaria `EDITORES_CD`. Los valores en la clave foránea pueden repetirse cualquier número de veces, a menos que se restrinja la columna por una restricción única.

Antes de pasar a discutir los otros elementos de la sintaxis de `FOREIGN KEY`, echemos un vistazo rápido a la clave foránea que incluye varias columnas. En la figura 4-6 hay dos tablas: `ARTISTAS_INTERPRETES` y `TIPOS_MUSICA_ARTISTAS`.

La clave primaria en la tabla `ARTISTAS_INTERPRETES` se define en las columnas `NOMBRE_ARTISTA` y `FDN_ARTISTA`. La siguiente instrucción de SQL crea la tabla `TIPOS_MUSICA_ARTISTAS`, que incluye una clave foránea formada por las columnas `NOMBRE_ARTISTA` y `FDN`:

```
CREATE TABLE TIPOS_MUSICA_ARTISTAS
( NOMBRE_ARTISTA VARCHAR(60),
  FDN DATE,
  ID_TIPO INT,
  CONSTRAINT FK_ARTISTAS_CD FOREIGN KEY ( NOMBRE_ARTISTA, FDN )
    REFERENCES ARTISTAS_INTERPRETES (NOMBRE_ARTISTA, FDN_ARTISTA) );
```



Figura 4-6 Una clave foránea formada por varias columnas.

En esta instrucción hay dos columnas de referencia (NOMBRE_ARTISTA y FDN) y dos columnas referenciadas (NOMBRE_ARTISTA y FDN_ARTISTA). Las columnas NOMBRE_ARTISTA en las dos tablas tienen el mismo tipo de datos, y la columna FDN tiene el mismo tipo de datos que la columna FDN_ARTISTA. Como se puede observar, una de las columnas de referencia (FDN) tiene un nombre diferente que su contraparte de referencia (FDN_ARTISTA).

Pregunta al experto

P: En la figura 4-6 y en los ejemplos anteriores, se creó una restricción FOREIGN KEY en las columnas NOMBRE_ARTISTA y FDN en la tabla TIPOS_MUSICA_ARTISTAS. ¿Cuál podría ser la clave primaria para esta tabla?

R: Recuerde que una clave primaria debe identificar de manera exclusiva cada fila en una tabla. Sin embargo, ya que los pares de valores en las columnas NOMBRE_ARTISTA y FDN se pueden repetir (lo que significa que también pueden repetirse en las columnas individuales), esas dos columnas no se pueden utilizar como una clave primaria para esta tabla. Por otro lado, la columna ID_TIPO también puede repetir valores; por lo tanto, esa columna no se puede utilizar. Además, probablemente no desee combinar la columna ID_TIPO con una de las

(continúa)

otras dos columnas porque es posible que se repitan filas (por ejemplo, dos artistas con el mismo nombre interpretando el mismo tipo de música, tal como los dos músicos de blues llamados Sonny Boy Williamson, o dos artistas con la misma fecha de nacimiento interpretando el mismo tipo de música). Como resultado, la mejor solución (además de añadir otra columna a la tabla) es mover las tres columnas a una clave primaria. Juntas, las tres columnas identifican de forma exclusiva cada fila, ya que es poco probable que alguien comparta el mismo nombre, fecha de nacimiento y tipo de música (aunque cualquier cosa es posible, razón por la cual, en última instancia, añadir otra columna que garantice que los datos son únicos es la mejor manera de hacerse).

La cláusula MATCH

Ahora que tiene una comprensión de cómo definir una restricción básica de clave foránea, echemos un vistazo de otra línea de sintaxis de FOREIGN KEY:

```
[ MATCH { FULL | PARTIAL | SIMPLE } ]
```

Se puede decir por los paréntesis que se trata de una cláusula opcional. Y de hecho, muy pocos proveedores de productos respaldan actualmente esta cláusula (no se respalda por SQL Server 2005, Oracle 11g o MySQL 5.0, por ejemplo), por lo que verá que no se utiliza mucho. Sin embargo, se describe en el estándar SQL, lo que significa que se espera que más proveedores de productos la respalden en el futuro. Su propósito es permitirle decidir cómo tratar los valores nulos en las columnas de clave foránea con respecto a los valores que permiten que se inserten en las columnas de referencia. Si las columnas no permiten valores nulos, entonces la cláusula MATCH no se aplica. Se tienen tres opciones que se pueden utilizar en la cláusula MATCH:

- Si se especifica MATCH FULL, todas las columnas de referencia deben tener un valor nulo o ninguna de esas columnas puede tener un valor nulo.
- Si se especifica MATCH PARTIAL y una o más columnas de referencia pueden tener valores nulos siempre y cuando el resto de las columnas de referencia tengan valores que igualen a las columnas de referencia correspondientes.
- Si se especifica MATCH SIMPLE y una o más columnas de referencia tienen valores nulos, entonces el resto de las columnas de referencia pueden tener valores que no están contenidos en las columnas de referencia correspondientes. La opción SIMPLE está implícita si no se incluye la cláusula MATCH en la definición de la restricción de FOREIGN KEY.

La mejor manera de ilustrar cada una de esas opciones MATCH es mediante ejemplos de datos válidos e inválidos que pueden insertarse en las columnas de referencia. Volviendo al ejemplo mostrado en la figura 4-6, se puede observar que una clave foránea en la tabla TIPOS_MUSICA_ARTISTAS se forma por dos columnas de referencia: NOMBRE_ARTISTA y FDN. La tabla 4-1 proporciona ejemplos de datos que se pueden o no insertar en las columnas de clave foránea. Los ejemplos se basan en los datos de las columnas de clave primaria de la tabla ARTISTAS_INTERPRETES.

Opción MATCH	Ejemplos de datos válidos	Ejemplos de datos inválidos
FULL	Joni Mitchell, 1943-11-07 NULL, NULL	NULL, 1943-11-07 Joni Mitchell, NULL Joni Mitchell, 1802-08-03
PARTIAL	Patsy Cline, 1932-09-08 NULL, 1932-09-08 Patsy Cline, NULL NULL, NULL	NULL, 1802-08-03 Henryk Górecki, NULL Patsy Cline, 1947-03-03
SIMPLE	Bing Crosby, 1904-05-02 NULL, 1904-05-02 Bing Crosby, NULL NULL, 1802-08-03 Henryk Górecki, NULL NULL, NULL	Bing Crosby, 1802-08-03 Bing Crosby, 1947-03-03 Henryk Górecki, 1947-03-03

Tabla 4-1 Ejemplos válidos e inválidos de las opciones de una cláusula MATCH.

NOTA

Probablemente no desea permitir que haya valores nulos en las columnas de referencia en la tabla TIPOS_MUSICA_ARTISTAS, particularmente en la columna NOMBRE_ARTISTA. Y si cualquiera de esas columnas se utiliza en la clave primaria, no sería posible permitir valores nulos. Sin embargo, con el fin de demostrar cómo trabajan las opciones MATCH, asumamos que se permiten los valores nulos.

Si decide utilizar una cláusula MATCH, simplemente añádala al final de la definición de la restricción FOREIGN KEY, como se muestra en la siguiente instrucción de SQL (asumiendo que la aplicación de SQL la respalda):

```
CREATE TABLE TIPOS_MUSICA_ARTISTAS
( NOMBRE_ARTISTA VARCHAR(60),
  FDN DATE,
  ID_TIPO INT,
  CONSTRAINT FK_ARTISTAS_CD FOREIGN KEY ( NOMBRE_ARTISTA, FDN )
    REFERENCES ARTISTAS_INTERPRETES MATCH FULL );
```

Para insertar datos en las columnas de referencia (NOMBRE_ARTISTA y FDN), ambos valores tienen que ser nulos o deben ser valores de datos válidos de las columnas referenciadas en la tabla ARTISTAS_INTERPRETES.

La cláusula <acción referencial desencadenada>

La cláusula final en la sintaxis de la restricción FOREIGN KEY es la cláusula opcional <acción referencial desencadenada>. La cláusula permite definir qué tipos de acciones se deben tomar cuando se intenta actualizar o eliminar datos desde columnas referenciadas (si ese intento causa una violación de los datos en las columnas de referencia). Por ejemplo, suponga que trata de eliminar datos de una tabla con clave primaria. Si esa clave primaria hace referencia a una clave foránea

y si los datos a eliminar se almacenan en una clave foránea, entonces eliminar los datos desde una clave primaria causaría una violación a la restricción FOREIGN KEY. Los datos en las columnas de referencia siempre deben incluirse en las columnas referenciadas.

El punto que debe recordarse acerca de la cláusula <acción referencial desencadenada> es que se incluye en la definición de la tabla de referencia (a través de una clave foránea) una acción que se debe tomar como resultado de algo que se realiza a la tabla referenciada. Esto se puede aclarar echando un vistazo a la sintaxis para la cláusula de <acción referencial desencadenada>:

```
ON UPDATE <acción referencial> [ ON DELETE <acción referencial> ]  
| ON DELETE <acción referencial> [ ON UPDATE <acción referencial> ]  
<acción referencial>::=  
CASCADE | SET NULL | SET DEFAULT | RESTRICT | NO ACTION
```

NOTA

El símbolo ::= (dos puntos consecutivos y un signo de igual) se utiliza en el estándar SQL:2006 para separar a un marcador de posición de las llaves de la definición. En la sintaxis anterior se define el marcador de posición <acción referencial>. El marcador de posición se utiliza en el código que precede a la definición. Se tomará entonces la definición (las cinco palabras clave) y se usarán en lugar del marcador de posición <acción referencial> como se utiliza en las cláusulas ON UPDATE y ON DELETE.

Como se observa de la sintaxis, se puede definir una cláusula ON UPDATE, y una cláusula ON DELETE, o ambas, y se pueden definir en cualquier orden. Para cada una de estas cláusulas se puede escoger una de cinco acciones referenciales:

- Si se utiliza CASCADE y los datos se actualizan o eliminan en las columnas referenciadas, los datos en las columnas de referencia se actualizan o eliminan.
- Si se utiliza SET NULL y los datos se actualizan o eliminan en las columnas referenciadas, los valores en las columnas de referencia correspondientes se establecen como nulos. Los valores nulos tienen que respaldarse en las columnas de referencia para que esta opción funcione.
- Si se utiliza SET DEFAULT y los datos se actualizan o eliminan en las columnas referenciadas, los valores en las columnas de referencia correspondientes se establecen a sus valores predeterminados. Se deben asignar valores predeterminados a las columnas de referencia para que esta opción funcione.
- Si se utiliza RESTRICT y se tratan de actualizar o eliminar datos en las columnas referenciadas que causan una violación de clave foránea, se previene la ejecución de esta acción. Los datos en las columnas de referencia nunca pueden violar la restricción FOREIGN KEY, ni siquiera temporalmente.
- Si se utiliza NO ACTION y se tratan de actualizar o eliminar datos en las columnas referenciadas que causan una violación de clave foránea, se previene la ejecución de esta acción. Sin embargo, las violaciones de datos pueden ocurrir temporalmente bajo ciertas condiciones durante la ejecución de una instrucción SQL, pero los datos en la clave foránea nunca se violan en su estado final (al final de la ejecución). La opción NO ACTION es predeterminada y se utiliza tanto para actualizar como para eliminar, si no se especifica una acción referencial desencadenada.

Si decide utilizar la cláusula <acción referencial desencadenada>, simplemente añádala al final de la definición de la restricción de FOREIGN KEY, como se muestra en la siguiente instrucción SQL:

```
CREATE TABLE TIPOS_MUSICA_ARTISTAS
( NOMBRE_ARTISTA VARCHAR(60),
  FDN DATE,
  ID_TIPO INT,
  CONSTRAINT FK_ARTISTAS_CD FOREIGN KEY ( NOMBRE_ARTISTA, FDN )
    REFERENCES ARTISTAS_INTERPRETES ON UPDATE CASCADE ON DELETE CASCADE );
```

Si actualiza o elimina datos de las columnas referenciadas en ARTISTAS_INTERPRETES, esos cambios se harán en las columnas de referencia en la tabla TIPOS_MUSICA_ARTISTAS.

Pruebe esto 4-1

Añada restricciones NOT NULL, únicas y referenciales

En el capítulo 3, Pruebe esto 3-1 y 3-2, se crearon varias tablas que se agregaron a la base de datos INVENTARIO (o el esquema CD_INVENTARIO). En este ejercicio se añadirán una serie de restricciones a las tablas y se crearán nuevas tablas que también se definirán con restricciones. Sin embargo, en lugar de utilizar la instrucción ALTER TABLE para modificar las tablas que ya se crearon, se volverán a crear esas tablas. La ventaja de esto es que podrá observar la definición de la tabla por completo en lo que se refiere a la actualización del modelo de datos, mostrado en la figura 4-7.

El modelo de datos incorpora un par de elementos más, que se han visto antes. Incorpora tablas, columnas dentro de esas tablas, tipos de datos para esas columnas, restricciones y relaciones entre tablas. Ya debe estar familiarizado con la forma en que se representan las tablas, columnas y tipos de datos, así que echemos un vistazo a las restricciones y relaciones:

- Las columnas incluidas en la clave primaria están en la sección superior de la tabla, y las otras columnas se encuentran en la sección inferior. Por ejemplo, en la tabla DISCOS_COMPACTOS, la columna ID_DISCO_COMPACTO es la clave primaria. En algunos casos, como en la tabla TIPOS_DISCOS_COMPACTOS, todas las columnas se incluyen en la clave primaria.
- Cada clave foránea se representa por [FK].
- Las predeterminaciones, las restricciones UNIQUE y las restricciones NOT NULL se identifican con cada columna aplicable.
- Las relaciones, tal como las claves foráneas las definen, se representan por líneas que conectan la clave foránea en una tabla a la clave de candidato (suele ser la clave primaria) en otra tabla.

Encontrará útil este modelo de datos no sólo en este ejercicio, sino para otros ejercicios en el libro, los cuales continúan aprovechando o utilizando la base de datos INVENTARIO. Puede también descargar el archivo Try_This_04.txt, que contiene las instrucciones SQL utilizadas en este ejercicio (en inglés).

(continúa)

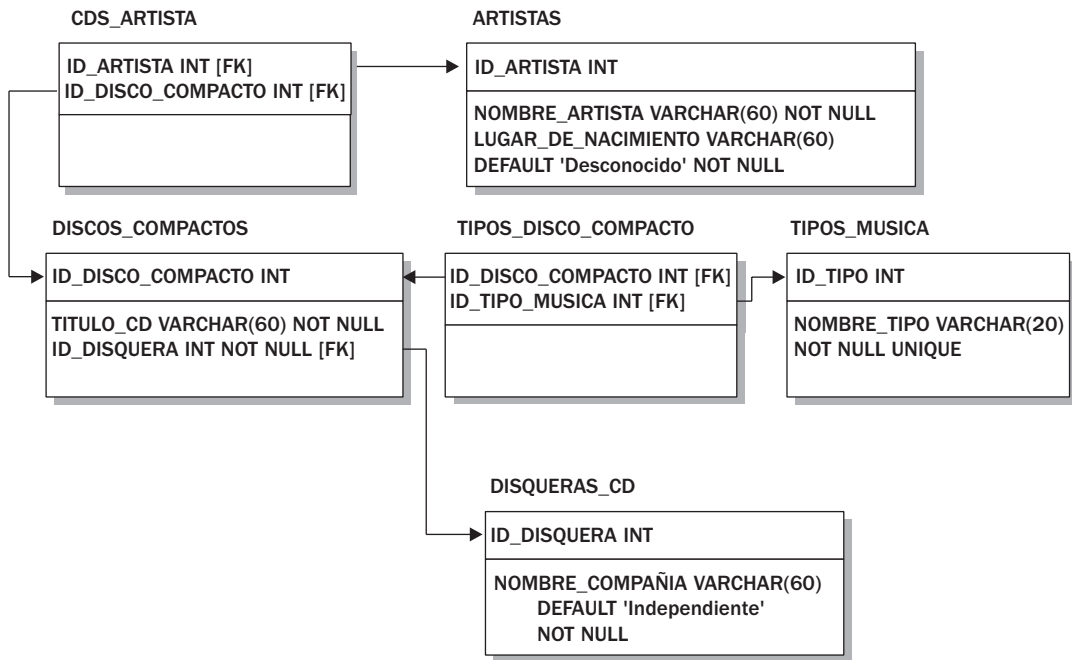


Figura 4-7 Modelo de datos para la base de datos INVENTARIO.

NOTA

Los modelos de datos son muy variados. El modelo que aquí se utiliza es específico para las necesidades del libro. Encontrará en el mundo real que los modelos son diferentes de lo que se ve aquí. Por ejemplo, las relaciones entre tablas se representan de manera diferente, y la información de la definición de columna podría no ser tan extensa.

Paso a paso

1. Abra la aplicación de clientes de su RDBMS y conecte la base de datos INVENTARIO.
2. Primero necesita eliminar las cuatro tablas (DISCOS_COMPACTOS, TIPOS_DISCO_COMPACTO, TIPOS_MUSICA y DISQUERAS_CD) que ya se crearon. Introduzca y ejecute las siguientes instrucciones:

```

DROP TABLE DISCOS_COMPACTOS      CASCADE;
DROP TABLE TIPOS_DISCO_COMPACTO  CASCADE;
DROP TABLE TIPOS_MUSICA          CASCADE;
DROP TABLE DISQUERAS_CD          CASCADE;
  
```


NOTA

Si creó la tabla ARTISTAS o la tabla CD_ARTISTA cuando probó con ejemplos o experimentó con las instrucciones CREATE TABLE, asegúrese de eliminar ésas también.

NOTA

SQL Server no soporta la opción CASCADE, y en Oracle se debe escribir como CASCADE CONSTRAINTS.

Ahora puede volver a crear estas tablas y crear otras nuevas. Debe crear las tablas en el orden indicado en este ejercicio, ya que las tablas referidas en las claves foráneas tienen que existir (con las claves primarias creadas) antes de que pueda crear las claves foráneas. Asegúrese de hacer referencia al modelo de datos de la figura 4-7 para obtener más detalles acerca de cada una de las tablas que se crean.

3. La primera tabla que se crea es la tabla TIPOS_MUSICA. Contiene dos columnas: ID_TIPO y NOMBRE_TIPO. Se configura la columna ID_TIPO como la clave primaria, y se configura la restricción UNIQUE y la restricción NOT NULL en la columna NOMBRE_TIPO. Introduzca y ejecute la siguiente instrucción SQL:

```
CREATE TABLE TIPOS_MUSICA
( ID_TIPO          INT,
  NOMBRE_TIPO      VARCHAR(20) NOT NULL,
  CONSTRAINT UN_NOMBRE_TIPO UNIQUE (NOMBRE_TIPO),
  CONSTRAINT PK_TIPOS_MUSICA PRIMARY KEY (ID_TIPO) );
```

4. La siguiente tabla que se crea es la tabla DISQUERAS_CD. La tabla incluye la columna ID_DISQUERA, que se define como la clave primaria, y la columna NOMBRE_COMPañIA, que se define con un valor predeterminado y una restricción NOT NULL. Introduzca y ejecute la siguiente instrucción SQL:

```
CREATE TABLE DISQUERAS_CD
( ID_DISQUERA      INT,
  NOMBRE_COMPañIA  VARCHAR(60) DEFAULT 'Independiente' NOT NULL,
  CONSTRAINT PK_DISQUERAS_CD PRIMARY KEY (ID_DISQUERA) );
```

5. Ahora que se creó la tabla DISQUERAS_CD, se puede crear la tabla DISCOS_COMPACTOS. La tabla DISCOS_COMPACTOS contiene una clave foránea que hace referencia a la tabla DISQUERAS_CD. Ésta es la razón por la que se creó primero la tabla DISQUERAS_CD. Introduzca y ejecute la siguiente instrucción SQL:

```
CREATE TABLE DISCOS_COMPACTOS
( ID_DISCO_COMPACTO INT,
  TITULO_CD          VARCHAR(60) NOT NULL,
  ID_DISQUERA        INT NOT NULL,
  CONSTRAINT PK_DISCOS_COMPACTOS PRIMARY KEY (ID_DISCO_COMPACTO),
  CONSTRAINT FK_ID_DISQUERA FOREIGN KEY (ID_DISQUERA) REFERENCES
DISQUERAS_CD );
```

(continúa)

6. La siguiente tabla, TIPOS_DISCO_COMPACTO, incluye dos claves foráneas, junto con su clave primaria. Las claves foráneas hacen referencia a las tablas DISCOS_COMPACTO y TIPOS_MUSICA, las cuales ya se habían creado. Introduzca y ejecute la siguiente instrucción SQL:

```
CREATE TABLE TIPOS_DISCO_COMPACTO
( ID_DISCO_COMPACTO INT,
  ID_TIPO_MUSICA INT,
  CONSTRAINT PK_TIPOS_DISCO_COMPACTO
    PRIMARY KEY ( ID_DISCO_COMPACTO, ID_TIPO_MUSICA),
  CONSTRAINT FK_ID_DISCO_COMPACTO_01
    FOREIGN KEY (ID_DISCO_COMPACTO) REFERENCES DISCOS_COMPACTOS,
  CONSTRAINT FK_ID_TIPO_MUSICA
    FOREIGN KEY (ID_TIPO_MUSICA) REFERENCES TIPOS_MUSICA );
```

7. Ahora ya se puede crear la tabla ARTISTAS. Introduzca y ejecute la siguiente instrucción SQL:

```
CREATE TABLE ARTISTAS
( ID_ARTISTA INT,
  NOMBRE_ARTISTA VARCHAR(60) NOT NULL,
  LUGAR_DE_NACIMIENTO VARCHAR(60) DEFAULT 'Desconocido' NOT NULL,
  CONSTRAINT PK_ARTISTAS PRIMARY KEY (ID_ARTISTA) );
```

8. La última tabla que se crea (al menos por ahora) es la tabla CDS_ARTISTA. Introduzca y ejecute la siguiente instrucción SQL:

```
CREATE TABLE CDS_ARTISTA
( ID_ARTISTA INT,
  ID_DISCO_COMPACTO INT,
  CONSTRAINT PK_CDS_ARTISTA PRIMARY KEY ( ID_ARTISTA, ID_DISCO_
    COMPACTO ),
  CONSTRAINT FK_ID_ARTISTA FOREIGN KEY (ID_ARTISTA) REFERENCES
    ARTISTAS,
  CONSTRAINT FK_ID_DISCO_COMPACTO_02 FOREIGN KEY (ID_DISCO_COMPACTO)
    REFERENCES DISCOS_COMPACTOS );
```

9. Cierre la aplicación de cliente.

Resumen de Pruebe esto

La base de datos cuenta ahora con seis tablas, cada una configurada con los valores predeterminados y las restricciones necesarias. En este ejercicio, se seguirá un orden específico para crear las tablas con el fin de aplicar más fácilmente las claves foráneas. Sin embargo, se pudieron haber creado las tablas en cualquier orden, sin las claves foráneas (a menos que las tablas referenciadas ya se hayan creado) y luego agregarse en la clave foránea, pero esto agrega pasos extra. De hecho, se hubieran podido modificar las tablas que existían antes de este ejercicio (en lugar de eliminarlas y después volver a crearlas), siempre y cuando se crearan las claves primarias (o restricciones UNIQUE) en las tablas referenciadas antes de crear las claves foráneas en las tablas de referencia. Independientemente del enfoque que tome, el resultado final debe ser que ahora la base de datos cuenta con las tablas necesarias para empezar a pasar a otros componentes de SQL.

Defina restricciones CHECK

Anteriormente en este capítulo, en la sección “Entienda las restricciones de integridad”, se analizaron varias categorías de restricciones y los tipos de restricciones que se soportan. (Remítase a la figura 4-1 para una descripción general de esas categorías.) Un tipo de restricción (la restricción CHECK) se puede definir como restricciones de tabla, restricciones de columna, restricciones de dominio, o en afirmaciones. Una restricción CHECK permite especificar qué valores se pueden incluir en una columna. Se puede definir un rango de valores (por ejemplo, entre 10 y 100), una lista de valores (por ejemplo, blues, jazz, pop, country), o una serie de otras condiciones que restringen exactamente qué valores se permiten en una columna.

Las restricciones CHECK son las más flexibles de todas las restricciones y suelen ser las más complicadas. A pesar de esto, la sintaxis básica para una restricción CHECK es relativamente sencilla. Para crear una restricción CHECK de columna, utilice la siguiente sintaxis en la definición de columna:

```
<nombre de columna> { <tipo de datos> | <dominio> } CHECK ( <condición de búsqueda> )
```

Para crear una restricción CHECK de tabla, utilice la siguiente sintaxis en la definición de la tabla:

```
[ CONSTRAINT <nombre de restricción> ] CHECK ( <condición de búsqueda> )
```

Se analizarán las restricciones de dominio y afirmaciones después en esta sección.

Como se puede observar por la sintaxis, una restricción CHECK es relativamente sencilla. Sin embargo, los valores utilizados para la cláusula <condición de búsqueda> pueden ser muy amplios y, por consiguiente, bastante complejos. El concepto principal es que la <condición de búsqueda> se pruebe (se puede decir “compruebe”) para cualquier instrucción SQL que intente modificar los datos en una columna cubierta por la restricción de CHECK, y si se evalúa como TRUE, la instrucción SQL se completa; si se evalúa como FALSE, la instrucción SQL se suspende y despliega un mensaje de error. La mejor manera de aprender acerca de la cláusula es examinar los ejemplos. Sin embargo, la mayoría de los componentes de la <condición de búsqueda> se basan en el uso de predicados con el fin de crear la condición de búsqueda. Un *predicado* es una expresión que opera en valores. Por ejemplo, un predicado se puede utilizar para comparar valores (por ejemplo, COLUMNA_1 > 10). El símbolo mayor que (>) es un predicado de comparación, a menudo denominado como *operador de comparación*. En este caso, el predicado verifica que cualquier valor insertado en la COLUMNA_1 sea mayor que 10.

Muchos de los componentes de la <condición de búsqueda> también se basan en el uso de subconsultas. Una *subconsulta* es una expresión que se utiliza como un componente dentro de otra expresión. Las subconsultas se usan cuando una expresión debe acceder o calcular varias capas de datos, tal como tener que buscar una segunda tabla para proporcionar datos para una primera tabla.

Tanto los predicados como las subconsultas son temas bastante complicados que van más allá del alcance de un análisis acerca de las restricciones CHECK, y de hecho, cada uno de los temas se tratan por separado en su propio capítulo. (Vea el capítulo 9 para información acerca de predicados y el capítulo 12 para información acerca de subconsultas.) A pesar del hecho de que ambos temas se analizan después en el libro, quiero proporcionar al menos unos ejemplos de las restricciones CHECK para dar una idea de cómo se implementan en el entorno SQL.

El primer ejemplo que se verá es una restricción CHECK que define valores mínimos y máximos que se pueden insertar en una columna. La siguiente definición de tabla en este ejemplo crea

tres columnas y una restricción CHECK (como una restricción de tabla) que restringe los valores de una de las columnas a una serie de números entre 0 y 30:

```
CREATE TABLE TITULOS_CD
( ID_DISCO_COMPACTO INT,
  TITULO_CD          VARCHAR(60) NOT NULL,
  EN_EXISTENCIA      INT          NOT NULL,
  CONSTRAINT CK_EN_EXISTENCIA CHECK ( EN_EXISTENCIA > 0 AND EN_
EXISTENCIA < 30 ) );
```

Si se trata de introducir un valor en la columna EN_EXISTENCIA distinto de 1 a 29, se recibe un error. Se pueden lograr los mismos resultados definiendo una restricción de columna:

```
CREATE TABLE TITULOS_CD
( ID_DISCO_COMPACTO INT,
  TITULO_CD          VARCHAR(60) NOT NULL,
  EN_EXISTENCIA      INT          NOT NULL
  CHECK ( EN_EXISTENCIA > 0 AND EN_EXISTENCIA < 30 ) );
```

Echemos un vistazo más de cerca a la cláusula <condición de búsqueda> en la instrucción, que en este caso es (EN_EXISTENCIA > 0 AND EN_EXISTENCIA < 30). La cláusula primero nos dice que cualquier valor introducido en la columna EN_EXISTENCIA debe ser mayor que 0 (EN_EXISTENCIA > 0). La palabra clave AND nos dice que las condiciones definidas en ambos lados de AND deben aplicarse. Finalmente, la cláusula nos dice que el valor debe ser menor que 30 (EN_EXISTENCIA < 30). Debido a que se usa la palabra clave AND, el valor debe ser mayor que 0 y menor que 30.

Otra manera en la que se puede utilizar una restricción CHECK es explícitamente listar los valores que se pueden introducir en la columna. Ésta es una opción práctica si se tiene un número limitado de valores y no son susceptibles de cambio (o cambian con poca frecuencia). La siguiente instrucción SQL crea una tabla que incluye una restricción CHECK que define a qué década pertenece la música:

```
CREATE TABLE TITULOS_CD
( ID_DISCO_COMPACTO INT,
  TITULO_CD          VARCHAR(60) NOT NULL,
  ERA                CHAR(5),
  CONSTRAINT CK_ERA CHECK ( ERA IN ( '1940s', '1950s',
    '1960s', '1970s', '1980s', '1990s', '2000s' ) ) );
```

El valor introducido en la columna ERA debe ser una de las siete décadas representadas en la condición de búsqueda. Si se trata de introducir un valor que no sea un valor nulo o uno de estos siete, se recibirá un error. Observe que se utiliza el operador IN para designar que los valores en la columna ERA deben ser uno del conjunto de valores delimitados entre paréntesis después de la palabra clave IN.

Si el número del paréntesis empieza a confundirlo, se puede separar el código en líneas que siguen a la incrustación de esos paréntesis. Por ejemplo, la instrucción anterior se puede escribir como sigue:

```
CREATE TABLE TITULOS_CD
(
```

```

ID_DISCO_COMPACTO    INT,
TITULO_CD             VARCHAR(60)  NOT NULL,
ERA                  CHAR(5),
CONSTRAINT CK_ERA     CHECK
(
    ERA IN
    (
        '1940s', '1950s', '1960s', '1970s', '1980s', '1990s', '2000s'
    )
)
);

```

Cada conjunto de paréntesis y su contenido se escriben a un nivel que corresponde al nivel de incrustación para esa cláusula en particular, al igual que un esquema. Utilizando este método nos dice exactamente qué cláusulas se incluyen en qué conjunto de paréntesis, y la instrucción se ejecuta de la misma forma como si no se hubiera separado las líneas. El inconveniente es que ocupa mucho espacio (que es la razón por la que no se utiliza este método en el libro), aunque podría ser una herramienta útil para esas instrucciones que son un poco más complejas.

Ahora veamos otro ejemplo de una restricción CHECK. Este ejemplo es similar al primero que examinamos, sólo que éste se refiere a valores entre determinados números:

```

CREATE TABLE TITULOS_CD
( ID_DISCO_COMPACTO    INT,
  TITULO_CD            VARCHAR(60)  NOT NULL,
  EN_EXISTENCIA        INT          NOT NULL,
  CONSTRAINT CK_EN_EXISTENCIA CHECK
    ( ( EN_EXISTENCIA BETWEEN 0 AND 30 ) OR
      ( EN_EXISTENCIA BETWEEN 49 AND 60 ) ) ) ;

```

En esta instrucción se utiliza el operador BETWEEN para especificar un rango que *incluye* los criterios de valoración. Ya que se crean dos diferentes rangos, se encierra la especificación de cada rango en paréntesis: (EN_EXISTENCIA BETWEEN 0 AND 30) y (EN_EXISTENCIA BETWEEN 49 AND 60). Estas dos especificaciones de rango se conectan por la palabra clave OR, lo que indica que una o la otra condición debe cumplirse. Como resultado, cualquier valor introducido en la columna EN_EXISTENCIA debe ser de 0 a 30 o de 49 a 60.

Como se mencionó antes, se aprenderá más acerca de condiciones de búsqueda en el capítulo 9. En ese momento se verá cuán flexible es la restricción CHECK. Y cuando se utiliza con subconsultas (vea capítulo 12), proporciona una herramienta poderosa para definir explícitamente qué valores se permiten en una columna en particular.

Defina afirmaciones

Una afirmación es simplemente un tipo de restricción CHECK que se puede aplicar a varias tablas. Por esta razón, una afirmación se crea por separado de la definición de una tabla. Desafortunadamente, la mayoría de los proveedores de productos, incluyendo Oracle 11g, SQL Server 2005 y MySQL 5.0, aún no respaldan las afirmaciones. Para crear una afirmación, utilice la siguiente sintaxis:

```
CREATE ASSERTION <nombre de la restricción> CHECK <condiciones de búsqueda>
```

Crear una afirmación es muy similar a crear una restricción CHECK. Después de la palabra clave CHECK, se debe proporcionar la(s) condición(es) de búsqueda necesaria(s). Ahora veamos un ejemplo. Suponga que la tabla TITULOS_CD incluye una columna para el número de discos compactos en existencia. Se desea que el total para esa tabla sea siempre inferior que el máximo de inventario que desea llevar. En el siguiente ejemplo, se crea una afirmación que totalice los valores en la columna EN_EXISTENCIA y verifique que el total sea menor que 5 000:

```
CREATE ASSERTION LIMITE_EN_EXISTENCIA CHECK
  ( ( SELECT SUM (EN_EXISTENCIA) FROM TITULOS_CD ) < 5000 );
```

En esta instrucción se utiliza la subconsulta (SELECT SUM (EN_EXISTENCIA) FROM TITULOS_CD) y se compara con 5 000. La subconsulta comienza con la palabra clave SELECT, que se usa para consultar datos en una tabla. La función SUM suma los valores en la columna EN_EXISTENCIA, y la palabra clave FROM especifica en qué tabla se encuentra la columna. Los resultados de esta subconsulta se comparan (utilizando el operador de comparación menor que) a 5 000. Si se intenta agregar un valor en la columna EN_EXISTENCIA que cause que el total sea superior a 5 000, se recibirá un error.

Creación de dominios y restricciones de dominio

El último tipo de restricción CHECK es del tipo que se inserta en una definición de dominio. En su mayor parte, la definición de restricción es similar a lo que ha visto antes, excepto que ésta no se vincula a una columna o tabla específica. De hecho, las restricciones de dominio utilizan la palabra clave VALUE cuando se refieren al valor en una columna definida con esa restricción en particular. Echemos un vistazo a la sintaxis para crear un dominio:

```
CREATE DOMAIN <nombre del dominio> [AS] <tipo de datos>
[DEFAULT <valor predeterminado>]
[CONSTRAINT <nombre de la restricción>] CHECK (<condición de búsqueda>)
```

Ya debe estar familiarizado con la mayoría de los elementos en esta sintaxis. Se discutieron los tipos de datos y las cláusulas predeterminadas en el capítulo 3, y la definición de la restricción es similar a lo que se ha visto hasta ahora en este capítulo.

En el siguiente ejemplo se crea un dominio que se basa en el tipo de datos INT y que requiere que todos los valores sean entre 0 y 30:

```
CREATE DOMAIN CANTIDAD_EN_EXISTENCIA AS INT
  CONSTRAINT CK_CANTIDAD_EN_EXISTENCIA CHECK (VALUE BETWEEN 0 AND 30);
```

El único elemento realmente nuevo aquí (distinto de la cláusula CREATE DOMAIN) es la palabra clave VALUE, la cual, como se dijo, se refiere al valor de la columna definida con el dominio CANTIDAD_EN_EXISTENCIA. Como resultado, si se trata de insertar un valor (en una de esas columnas) que no sea entre 0 y 30, se recibirá un error.

Pruebe esto 4-2 Añada una restricción CHECK

En este ejercicio, que es relativamente corto, se utiliza la instrucción `ALTER TABLE` para modificar la tabla `DISCOS_COMPACTOS`. Se añade una columna a la tabla y luego se define una restricción `CHECK` que restrinja los valores que se pueden introducir en la columna. La columna adicional y la restricción no tienen un impacto en las otras tablas de la base de datos `INVENTARIO` o en la relación entre las tablas. Puede descargar el archivo `Try_This_04.txt` (en inglés), que contiene las instrucciones SQL que se usan en este ejercicio.

Paso a paso

1. Abra la aplicación de clientes de su RDBMS y conéctese a la base de datos `INVENTARIO`.
2. Se modifica la tabla `DISCOS_COMPACTOS` añadiendo la columna `EN_EXISTENCIA`. Introduzca y ejecute la siguiente instrucción SQL:

```
ALTER TABLE DISCOS_COMPACTOS
ADD COLUMN EN_EXISTENCIA INT NOT NULL;
```

NOTA

Para Oracle y SQL Server, omita la palabra clave `COLUMN`.

3. Ahora que la columna existe, se puede agregar una restricción `CHECK` a la definición de la tabla. Se puede introducir la restricción como una restricción de columna, pero añadiéndola por separado como una restricción de tabla le permitirá hacer cada paso por separado para que pueda ver los resultados de sus acciones. La restricción `CHECK` limita los valores que se pueden introducir en la columna `EN_EXISTENCIA`. Cada valor debe ser mayor que 0, pero menor que 50. Introduzca y ejecute la siguiente instrucción SQL:

```
ALTER TABLE DISCOS_COMPACTOS
ADD CONSTRAINT CK_EN_EXISTENCIA CHECK ( EN_EXISTENCIA > 0 AND
EN_EXISTENCIA < 50 );
```

4. Cierre la aplicación de cliente.

Resumen de Pruebe esto

La nueva columna, `EN_EXISTENCIA`, sigue el número de cada disco compacto que figura en la tabla `DISCOS_COMPACTOS`. La restricción `CK_EN_EXISTENCIA` restringe el número por fila a una cantidad entre 0 y 50. Ahora que la tabla se ha actualizado, no se puede añadir ningún valor que viole la restricción.

Autoexamen Capítulo 4

1. ¿Cuáles son las tres categorías de las restricciones de integridad?
2. ¿Cuáles son las diferencias entre una restricción de columna y una restricción de tabla?
3. ¿Qué tipos de restricciones se pueden incluir en una definición de columna?
4. ¿Cuál es la diferencia entre una restricción de tabla y una afirmación?
5. ¿Qué significa un valor nulo?
6. ¿Cuál de los siguientes tipos de restricciones soporta restricciones NOT NULL?
 - A Restricciones de tabla
 - B Restricciones de columna
 - C Restricciones de dominio
 - D Afirmaciones
7. Se crea una tabla que incluye una columna que acepta valores nulos pero cuyos valores no nulos deben ser únicos. ¿Qué tipo de restricción se debe utilizar?
8. Se crea una tabla que incluye la columna NOMBRE_TIPO. La columna se define con el tipo de datos CHAR(10) y requiere una restricción UNIQUE, que se define como una restricción de columna. ¿Qué código SQL se debe utilizar para la definición de columna?
9. ¿Cuáles dos restricciones se aplican a las restricciones PRIMARY KEY pero no aplican a las restricciones UNIQUE?
10. Se crea una restricción de PRIMARY KEY llamada PK_TIPOS_MUSICA_ARTISTA en la tabla TIPOS_MUSICA_ARTISTA. La clave primaria incluye las columnas NOMBRE_ARTISTA y FDN_ARTISTA. ¿Qué código SQL se debe utilizar para la restricción de la tabla?
11. ¿Cómo difiere una restricción referencial de una restricción única?
12. Una restricción _____ impone la integridad referencial entre dos tablas garantizando que no se lleve a cabo ninguna acción en ninguna tabla que pueda afectar a los datos protegidos por la restricción.
13. Se crea una tabla que incluye la columna llamada ID_TIPO_NEGOCIO, con un tipo de datos INT. La columna se define con una restricción FOREIGN KEY que hace referencia a la clave primaria en la tabla llamada TIPOS_NEGOCIO. La clave foránea se añade como una restricción de columna. ¿Qué código SQL se debe usar para la definición de columna?
14. ¿Cuáles tres opciones se pueden utilizar en la cláusula MATCH de una restricción FOREIGN KEY?
15. ¿Cuáles son los dos tipos de acciones referenciales desencadenadas que se pueden definir en una restricción FOREIGN KEY?

- 16.** Se crea una restricción FOREIGN KEY y se desea que los valores en la columna de referencia se actualicen si los valores en la columna referenciada se actualizan. ¿Qué cláusula <acción referencial desencadenada> utilizaría?
- A** ON UPDATE RESTRICT
 - B** ON UPDATE NO ACTION
 - C** ON UPDATE CASCADE
 - D** ON UPDATE SET DEFAULT
- 17.** ¿Qué sintaxis debe utilizar para una restricción CHECK que se define como una restricción de tabla?
- 18.** ¿Qué tipo de restricciones se pueden definir dentro en una afirmación?
- 19.** Se crea una restricción CHECK en la columna NUMERO_EN_EXISTENCIA. Se desea limitar los valores que se pueden introducir en la columna en un rango de 11 a 29. ¿Qué debe utilizar para la cláusula <condición de búsqueda> de la restricción?



Capítulo 5

Creación de vistas en SQL

Habilidades y conceptos clave

- Añada vistas a la base de datos
 - Creación de vistas actualizables
 - Eliminación de vistas de la base de datos
-

Como aprendió en el capítulo 3, las tablas base persistentes almacenan los datos de SQL en su base de datos. Sin embargo, esas tablas no son siempre una forma útil si sólo desea ver datos específicos de una tabla o datos de varias tablas. Por esta razón, el estándar SQL:2006 respalda el uso de tablas vistas, o vistas. Una *vista* es una tabla virtual cuya definición existe como un objeto de esquema. A diferencia de las tablas base persistentes, en la vista no hay datos almacenados. De hecho, las tablas vistas en realidad no existen (sólo existe la definición que las define). Esta definición es la que permite seleccionar información específica de una o más tablas, basada en las instrucciones de consulta en esa definición. Una vez que se crea una vista, simplemente se invoca llamándola por su nombre en una consulta como en una tabla base. Los datos entonces se presentan como si se buscaran en una tabla base.

Añada vistas a la base de datos

Antes de entrar demasiado profundo en las características específicas de las vistas, quiero revisar rápidamente algunos puntos que se discutieron en los capítulos 2 y 3. Una vista, como posiblemente recuerde, es uno de los tres tipos de tablas respaldadas por SQL, junto con las tablas base y las tablas derivadas. La mayoría de las tablas base son objetos de esquema y vienen en cuatro tipos: tablas base persistentes, tablas temporales globales, tablas temporales locales creadas y tablas temporales locales declaradas. De estos cuatro tipos, las tablas base persistentes son las que manejan los datos reales de SQL. Las tablas derivadas, por otro lado, son simplemente el resultado que se observa cuando se consultan datos de una base de datos. Por ejemplo, si se solicitan datos de la tabla DISCOS_COMPACTOS, los resultados de la solicitud se despliegan en formato parecido a una tabla, conocida como tabla derivada.

En algunos aspectos, una vista se encuentra entre una tabla base persistente y una tabla derivada. Es como una tabla base persistente, ya que la definición de la vista se almacena como un objeto de esquema utilizando un nombre único (dentro del esquema) que se puede acceder como en una tabla base. Sin embargo, una vista es como una tabla derivada, ya que ningún dato se almacena en asociación con la vista. Tanto las tablas derivadas como las vistas son tipos de tablas virtuales. Los datos se seleccionan de una o más tablas base cuando se invoca la vista. De hecho, se puede pensar en una vista como simplemente una tabla derivada nombrada, con la definición de la vista almacenada en el esquema. Los resultados de los datos que se ven cuando se llama a una vista se almacenan en cualquier lugar, pero se derivan de tablas base existentes.

Las vistas pueden ser herramientas útiles cuando se accede a los diferentes tipos de datos. Una de las principales ventajas de utilizar las vistas es que se pueden definir consultas complejas

y almacenarlas dentro de la definición de la vista. En lugar de volver a crear las consultas cada vez que se necesiten, puede simplemente invocar la vista. Además, las vistas pueden ser una manera práctica de presentar información a los usuarios sin suministrar más información que la que necesitan o información que no deben ver. Por ejemplo, puede ser que desee que los usuarios de su organización puedan acceder al registro de determinados empleados, pero puede que no quiera que la información tal como los números de seguro social o los sueldos estén disponibles para esos usuarios, así que se puede crear una vista que proporcione sólo la información que los usuarios deben ver. Las vistas también se pueden utilizar para sintetizar estructuras complejas y presentar información de una manera que sea más fácil de comprender para algunos usuarios, que en efecto oculta la estructura fundamental y la complejidad de la base de datos de los usuarios.

Ahora que tiene una descripción general de qué son las vistas, veamos algunos ejemplos que ilustren cómo se extraen los datos de tablas base en el tipo de tabla derivada que se presenta por la definición de la vista. El primer ejemplo que veremos, mostrado en la figura 5-1, se basa en la tabla `INVENTARIO_DISCO_COMPACTO`, que incluye seis columnas. Suponga que desea ver sólo las columnas `TITULO_CD`, `DERECHOSDEAUTOR` y `EN_EXISTENCIA`. Se puede crear una vista que extraiga estas tres columnas de la tabla y las organice como si los datos existieran en su propia tabla, como se muestra en la figura 5-1. La vista `DISCOS_COMPACTOS_EN_EXISTENCIA` contiene una consulta que define exactamente qué datos debe devolver la vista.

Se puede observar que los nombres de las columnas en la vista son diferentes de los nombres de las columnas de la tabla `INVENTARIO_DISCO_COMPACTO`, aun cuando los datos dentro de las columnas son los mismos. Esto se debe a que, si se desea, se puede asignar que los nombres de las columnas de las vistas sean diferentes a los de la tabla original. Lo mismo ocurre con los tipos de datos. Las columnas de la vista heredan los tipos de datos a partir de sus respectivas columnas de la tabla. Por ejemplo, la columna `DISCO_COMPACTO` en la vista `DISCOS_COMPACTOS_EN_EXISTENCIA` heredó el tipo de datos `VARCHAR(60)` de la columna `TITULO_CD` de la tabla `INVENTARIO_DISCO_COMPACTO`. No se especificó el tipo de datos `VARCHAR(60)` en ningún lugar dentro de la definición de la vista.

Como se puede observar, una vista permite definir qué columnas se devuelven cuando se invoca la vista. La definición para la vista `DISCOS_COMPACTOS_EN_EXISTENCIA` especifica tres columnas; sin embargo, se pudo haber especificado cualquiera de las columnas de la tabla `INVENTARIO_DISCO_COMPACTO`. Además de las columnas, la definición de una vista puede especificar qué filas se devuelven. Por ejemplo, la figura 5-2 muestra la vista `CDS_EN_EXISTENCIA_1990S`. Observe que contiene las mismas columnas que la vista `DISCOS_COMPACTOS_EN_EXISTENCIA` (mostrada en la figura 5-1), pero tiene menos filas. En este caso, la definición de la vista no sólo especifica las mismas tres columnas de la tabla `INVENTARIO_DISCO_COMPACTO`, sino que también especifica que sólo las filas con valores entre 1990 y 1999 (inclusive) en la columna `DERECHOSDEAUTOR` se devuelven.

En los dos ejemplos anteriores se examinaron las vistas que se derivan de datos de sólo una tabla; sin embargo, se pueden crear vistas basadas en varias tablas. Esto es particularmente útil si se quiere desplegar información relacionada que abarque más de una tabla. Echemos un vistazo a la figura 5-3, que incluye las tablas `INVENTARIO_CD` y `DISQUERAS`. La tabla `INVENTARIO_CD` contiene un listado de los CD en el inventario, y la tabla `DISQUERAS` contiene un listado de las compañías que editan CD. Se dará aquí una breve introducción del acceso a múltiples tablas, ya que es una gran manera de demostrar qué tan bien pueden las vistas ocultar la complejidad de una consulta. Este tema se trata en detalle en el capítulo 11.

INVENTARIO_DISCO_COMPACTO

ID_DISCO_COMPACTO: INT	TITULO_CD: VARCHAR(60)	COPYRIGHT: INT	ID_DISQUERA: INT	ID_DISCO: INT	EN_EXISTENCIA: INT
99301	Famous Blue Raincoat	1991	5422	1299	6
99302	Blue	1971	5402	1232	26
99303	Court and Spark	1974	5270	1287	18
99304	Past Light	1983	5412	1292	2
99305	Kojiki	1990	5409	1255	5
99306	That Christmas Feeling	1993	5403	1216	3
99307	Patsy Cline: 12 Greatest Hits	1988	5403	1210	25
99308	Carreras Domingo Pavarotti in Concert	1990	5312	1276	22
99310	Henryk Gorecki: Symphony No. 3	1992	5270	1266	8

DISCOS_COMPACTOS_
EN_EXISTENCIA

DISCO_COMPACTO	COPYRIGHT	EN_EXISTENCIA
Famous Blue Raincoat	1991	6
Blue	1971	26
Court and Spark	1974	18
Past Light	1983	2
Kojiki	1990	5
That Christmas Feeling	1993	3
Patsy Cline: 12 Greatest Hits	1988	25
Carreras Domingo Pavarotti in Concert	1990	22
Henryk Gorecki: Symphony No. 3	1992	8

Figura 5-1 La vista DISCOS_COMPACTOS_EN_EXISTENCIA, basada en la tabla INVENTARIO_DISCO_COMPACTO.

Suponga que tiene usuarios que desean poder ver los nombres del CD y del editor, pero que no les interesan los valores ID_DISCO_COMPACTO e ID_DISQUERA. Y ciertamente no les interesa tener que ver en dos lugares diferentes para comparar los valores ID_DISQUERA con el fin de igualar los CD con los nombres de la compañía. Una solución es crear una vista que iguale esta información, mientras que al mismo tiempo se despliegue sólo la información que les sea útil.

CDS_EN_EXISTENCIA_1990S

DISCO_COMPACTO	COPYRIGHT	EN_EXISTENCIA
Famous Blue Raincoat	1991	6
Kojiki	1990	5
That Christmas Feeling	1993	3
Carreras Domingo Pavarotti in Concert	1990	22
Henryk Gorecki: Symphony No. 3	1992	8

Figura 5-2 La vista CDS_EN_EXISTENCIA_1990S, basada en la tabla INVENTARIO_DISCO_COMPACTO.

INVENTARIO_CD

ID_DISCO_COMPACTO: INT	TITULO_CD: VARCHAR(60)	ID_DISQUERA: INT
11001	Famous Blue Raincoat	5422
11002	Blue	5402
11003	Past Light	5412
11004	Kojiki	5409
11005	That Christmas Feeling	5403
11006	Patsy Cline: 12 Greatest Hits	5403

DISQUERAS

ID_DISQUERA: INT	NOMBRE_COMPAÑIA: VARCHAR(60)
5403	MCA Records
5402	Reprise Records
5409	Geffen
5412	Windham Hill Records
5422	Private Music

EDITORES_DISCOS_COMPACTOS

DISCO_COMPACTO	EDITOR
Famous Blue Raincoat	Private Music
Blue	Reprise Records
Past Light	Windham Hill Records
Kojiki	Geffen
That Christmas Feeling	MCA Records
Patsy Cline: 12 Greatest Hits	MCA Records

Figura 5-3 La vista EDITORES_DISCOS_COMPACTOS, basada en las tablas DISQUERAS e INVENTARIO_CD.

En el caso de las tablas INVENTARIO_CD y DISQUERAS, se puede crear una vista (llamada EDITORES_DISCO_COMPACTO en la figura 5-3) que vincule (*una*) los datos para los usuarios, al tiempo que oculte la estructura fundamental y datos extraños.

Una vista de este tipo es posible que tome ventaja de las relaciones entre tablas. En el caso de las tablas INVENTARIO_CD y DISQUERAS, una clave foránea se definió en la columna ID_DISQUERA de la tabla INVENTARIO_CD que hace referencia a la columna ID_DISQUERA de la tabla DISQUERAS. La consulta contenida en la definición de la vista EDITORES_DISCO_COMPACTO coincide con los valores en la columna ID_DISQUERA de la tabla INVENTARIO_CD con los valores en la columna ID_DISQUERA de la tabla DISQUERAS. Para cada coincidencia que se encuentra, se devuelve una fila. Por ejemplo, la fila Famous Blue Raincoat incluye el valor ID_DISQUERA de 5422. En la tabla DISQUERAS se puede ver que ese valor coincide con la fila Private Music. Como resultado, la vista contiene una fila con los valores Famous Blue Raincoat y Private Music.

NOTA

No necesariamente se tiene que utilizar una relación de clave foránea para unir las tablas. Se pueden utilizar dos columnas de diferentes tablas que almacenen la misma información. Esto podría significar utilizar todas las columnas en una clave foránea (si la clave foránea incluye múltiples columnas), usando sólo una de las columnas, o sin utilizar una clave foránea. Se analizará la unión de múltiples tablas en el capítulo 11.

Además de unir información de tablas diferentes, también se pueden utilizar las vistas para modificar los datos que se extraen de la columna de una tabla y se presentan en una columna de vista. Esto permite tomar acciones como la realización de cálculos, encontrar promedios, determinar valores mínimos y máximos, y completar un sinnúmero de otras operaciones. Entonces se pueden tomar los resultados de estas operaciones e incluirlas en una columna dentro de una vista. En la figura 5-4, por ejemplo, la vista DESCUENTOS_CD deduce un 10 por ciento de descuento del precio al menudeo y presenta el resultado en la columna PRECIO_DESCUENTO.

La vista DESCUENTOS_CD incluye tres columnas. La columna DISCO_COMPACTO extrae los datos directamente de la columna TITULO_CD. Las columnas PRECIO_MENUDEO y PRECIO_DESCUENTO en la vista extraen los datos de la columna PRECIO_MENUDEO en la tabla INVENTARIO. La columna PRECIO_MENUDEO en la vista copia los valores justo como son. Sin embargo, para la columna PRECIO_DESCUENTO, los valores extraídos de la columna PRECIO_MENUDEO en la tabla INVENTARIO se multiplican por 0.9.

Como puede observar, se pueden especificar varios tipos de operaciones en una vista y luego simplemente invocar la vista cuando necesite la información. La mayor parte de lo que puede incluirse en una consulta se puede incluir en una vista. De hecho, es la consulta, o la expresión de consulta, la que constituye el núcleo de la vista. Sin embargo, antes de examinar las expresiones de consulta, deseo en primer lugar analizar la sintaxis utilizada para la creación de las vistas.

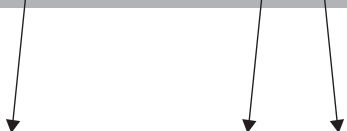
Definición de vistas de SQL

El tipo más simple de vista a crear es uno que hace referencia sólo a una tabla y recupera los datos de columnas dentro de una tabla sin modificar esos datos. Entre más complicada es la vista, más

INVENTARIO

ID_DISCO_COMPACTO: INT	TITULO_CD: VARCHAR(60)	COPYRIGHT: INT	PRECIO_MENUEDO: NUMÉRICO(5,2)	EN_EXISTENCIA: INT
99301	Famous Blue Raincoat	1991	16.99	6
99302	Blue	1971	14.99	26
99303	Court and Spark	1974	14.99	18
99304	Past Light	1983	15.99	2
99305	Kojiki	1990	15.99	5
99306	That Christmas Feeling	1993	10.99	3
99307	Patsy Cline: 12 Greatest Hits	1988	16.99	25

DESCUENTOS_CD



DISCO_COMPACTO	PRECIO_MENUEDO	PRECIO_DESCUENTO
Famous Blue Raincoat	16.99	15.29
Blue	14.99	13.49
Court and Spark	14.99	13.49
Past Light	15.99	14.39
Kojiki	15.99	14.39
That Christmas Feeling	10.99	9.89
Patsy Cline: 12 Greatest Hits	16.99	15.29

Figura 5-4 La vista DESCUENTOS_CD, basada en la tabla INVENTARIO.

complicada es la expresión de consulta en la que se basa la vista. En su forma más básica, la sintaxis para una vista es la siguiente:

```
CREATE VIEW <nombre de la vista> [(<nombres de las columnas de la vista>)]
AS <expresión de consulta>
[ WITH CHECK OPTION ]
```

Por ahora nos centraremos sólo en las primeras dos líneas de la sintaxis y se dejará WITH CHECK OPTION para después, en la sección “Creación de vistas actualizables”. Como se mues-

tra en la primera línea de la sintaxis, se debe proporcionar un nombre para la vista. Además, se deben proporcionar los nombres para las columnas en las siguientes circunstancias:

- Si alguno de los valores de columna se basa en algún tipo de operación que calcule el valor que se insertará en la columna, en lugar de que el valor se copie directamente de la tabla. (Ver la figura 5-4.)
- Si los nombres de las columnas de la tabla se duplican, lo cual puede suceder cuando las tablas se unen.

Incluso si no se requiere proporcionar los nombres de las columnas, se pueden proporcionar si así lo desea. Por ejemplo, puede encontrar que desea cambiar los nombres para que sean más lógicos para usuarios en particular. Sin embargo, si proporciona los nombres de las columnas utilizando la sintaxis <nombres de las columnas de la vista>, se deben proporcionar los nombres para todas las columnas.

NOTA

También hay una forma alternativa de proporcionar los nombres de las columnas utilizando la palabra clave AS dentro de la expresión de consulta, que se analizará en el capítulo 7.

La segunda línea de la sintaxis incluye la palabra clave AS, que es requerida, y el marcador de posición <expresión de consulta>. El marcador de posición <expresión de consulta>, aunque parece sencillo, puede implicar una estructura compleja de instrucciones de consulta que pueden realizar una serie de operaciones, incluyendo la recuperación de datos de varias tablas, el cálculo de datos, limitar el tipo de datos devueltos, y la realización virtual de cualquier otro tipo de operación respaldada por una expresión de consulta. Debido a la complejidad de las expresiones de consulta, pasaré la mayoría de la segunda parte de este libro discutiendo diversas formas de consulta de datos. Lo que implica, entonces, que será muy difícil resumir un debate a fondo de las expresiones de consulta dentro del tema de las vistas. A pesar de todo, quiero proporcionar una serie de ejemplos que ilustren cómo se pueden crear vistas que realicen varias funciones. Con cada ejemplo, se incluirá una breve explicación de la expresión de consulta utilizada en la definición de la vista. Esté al tanto, sin embargo, que se tratarán los detalles de las expresiones de consulta con mayor profundidad, más adelante en el libro, empezando en el capítulo 7. También observe que las tablas base utilizadas en estos ejemplos no se han creado en los ejercicios anteriores; por lo tanto, si quiere probar los ejemplos, se tendrán que crear primero las tablas base. Las figuras muestran la información que necesita para hacer esto.

El primer ejemplo que se considera se basa en la vista mostrada en la figura 5-1. La vista DISCOS_COMPACTOS_EN_EXISTENCIA obtiene los datos de la tabla INVENTARIO_DISCO_COMPACTO e incluye tres columnas de esa tabla. Para crear la vista, utilice la siguiente instrucción CREATE VIEW:

```
CREATE VIEW DISCOS_COMPACTOS_EN_EXISTENCIA
( DISCO_COMPACTO, DERECHOSDEAUTOR, EN_EXISTENCIA ) AS
  SELECT TITULO_CD, DERECHOSDEAUTOR, EN_EXISTENCIA
  FROM INVENTARIO_DISCO_COMPACTO;
```

Esta vista es la más simple de todos los tipos de vistas a crear. Se basa en una tabla y extrae tres de las seis columnas de esa tabla. Tenga en cuenta que mientras SQL normalmente requiere

cláusulas para estar en un orden particular, no hay restricciones en cuanto a espacios y saltos de línea. Por ejemplo, cuando se crean las vistas, se prefiere poner la lista de nombres de columna (cuando se presente) en una línea nueva y colocar la palabra clave AS al final de la línea que precede la expresión de consulta. Otros prefieren colocar la palabra clave AS en una línea, y aun otros prefieren colocarla al comienzo de la primera línea de la expresión de consulta. El RDBMS no tomará en cuenta la forma en que lo hace, pero la adopción de un estilo y el adherirse a él hará que SQL sea más fácil de leer, entender y mantener.

Analizando la instrucción un poco, la primera línea proporciona un nombre para la vista, DISCOS_COMPACTOS_EN_EXISTENCIA. La segunda línea proporciona un nombre para cada una de las tres columnas: DISCO_COMPACTO, DERECHOSDEAUTOR y EN_EXISTENCIA, y termina con la palabra clave AS. Si se omitieran los nombres de las columnas, las columnas de la vista tomarían los nombres de las columnas de la tabla. La tercera y cuarta líneas de la instrucción CREATE VIEW contienen la expresión de consulta, que en este caso es la siguiente instrucción SELECT:

```
SELECT DISCO_COMPACTO, DERECHOSDEAUTOR, EN_EXISTENCIA
FROM INVENTARIO_DISCO_COMPACTO
```

La instrucción SELECT es una de las instrucciones más comunes (si no *la* instrucción más común) que utilizará como programador de SQL. También es una de las instrucciones más extensas y flexibles que usará, permitiendo formar consultas intrincadas que pueden devolver exactamente el tipo de datos que se desee recuperar de la base de datos.

La instrucción SELECT utilizada en la definición de la vista DISCOS_COMPACTOS_EN_EXISTENCIA es la instrucción SELECT en su forma más básica. La instrucción se divide en dos cláusulas: la cláusula SELECT y la cláusula FROM. La cláusula SELECT identifica qué columnas se devuelven (TITULO_CD, DERECHOSDEAUTOR y EN_EXISTENCIA), y la cláusula FROM identifica la tabla de la que se extraen los datos (INVENTARIO_DISCO_COMPACTO). Cuando se invoca la vista DISCOS_COMPACTOS_EN_EXISTENCIA, se está esencialmente invocando la instrucción SELECT que se incrusta en la definición de la vista, que a su vez toma los datos de la(s) tabla(s) base aplicable(s).

En el siguiente ejemplo, basado en la vista de la figura 5-2, la instrucción CREATE VIEW es casi la misma que la del ejemplo anterior, excepto que una cláusula adicional se agrega a la instrucción:

```
CREATE VIEW CDS_EN_EXISTENCIA_1990S
( DISCO_COMPACTO, DERECHOSDEAUTOR, EN_EXISTENCIA ) AS
SELECT TITULO_CD, DERECHOSDEAUTOR, EN_EXISTENCIA
FROM INVENTARIO_DISCO_COMPACTO
WHERE DERECHOSDEAUTOR > 1989 Y DERECHOSDEAUTOR < 2000;
```

La cláusula WHERE define una condición que debe cumplirse para que los datos sean devueltos. Como en el ejemplo anterior, se extraen los datos de las columnas TITULO_CD, DERECHOSDEAUTOR y EN_EXISTENCIA de la tabla INVENTARIO_DISCO_COMPACTO, sólo que esta vez se limitan los datos a las filas cuyos valores DERECHOSDEAUTOR son mayores que 1989 pero menores que 2000 (DERECHOSDEAUTOR > 1989 Y DERECHOSDEAUTOR < 2000). Puede que reconozca los operadores de comparación mayor que (>) y menor que (<) del capítulo 4 del análisis acerca de las restricciones CHECK. Se utilizan para limitar los valores que se incluyen en la vista.

NOTA

Los operadores utilizados en la cláusula WHERE (o cualquier condición definida en la cláusula) no tienen efecto en los datos almacenados en las tablas base. Sólo afectan a los datos devueltos cuando se invoca la vista. Se analizarán estos tipos de operadores a mayor detalle en el capítulo 9.

Se puede utilizar la cláusula WHERE en la instrucción SELECT para definir una amplia variedad de condiciones. Por ejemplo, la cláusula WHERE se puede usar para ayudar a unir las tablas, como se muestra en la siguiente instrucción CREATE VIEW:

```
CREATE VIEW EDITORES_DISCO_COMPACTO
( DISCO_COMPACTO, EDITOR ) AS
SELECT INVENTARIO_CD.TITULO_CD, DISQUERAS.NOMBRE_COMPANIA
FROM INVENTARIO_CD, DISQUERAS
WHERE INVENTARIO_CD.ID_DISQUERA = DISQUERAS.ID_DISQUERA;
```

Esta instrucción crea la vista que se observa en la figura 5-3. El nombre de la vista es EDITORES_DISCO_COMPACTO e incluye las columnas DISCO_COMPACTO y EDITOR. La vista extrae información de dos fuentes: la columna TITULO_CD en la tabla INVENTARIO_CD y la columna NOMBRE_COMPANIA en la tabla DISQUERAS.

Primero echemos un vistazo a la cláusula SELECT. Observe que el nombre de cada columna se califica por el nombre de la tabla respectiva (por ejemplo, INVENTARIO_CD.TITULO_CD). Cuando se unen dos o más tablas, se deben calificar los nombres de las columnas si hay alguna posibilidad de que los nombres de las columnas se puedan confundir, que sería el caso si se incluyen columnas con el mismo nombre. Sin embargo, si no hay posibilidad de que los nombres de las columnas se puedan confundir, entonces se pueden omitir los nombres de la tabla. Por ejemplo, la cláusula SELECT puede leerse como se muestra a continuación:

```
SELECT TITULO_CD, NOMBRE_COMPANIA
```

A pesar de que los nombres calificados no siempre son necesarios, muchos programadores prefieren utilizarlos en todos los casos, ya que es más fácil saber a qué tabla se hace referencia si alguna vez necesita modificar la estructura de la base de datos o la definición de la vista en un momento posterior.

La siguiente cláusula en la instrucción SELECT es la cláusula FROM. Cuando se unen las tablas, se deben incluir los nombres de todas las tablas participantes, separadas por comas. Además de la cuestión de los diversos nombres, la cláusula FROM es similar a lo que se ha visto en otros ejemplos.

La cláusula WHERE, que es la cláusula final en la instrucción SELECT, es la que une las filas. La cláusula WHERE es necesaria, ya que sin ella no habría forma de saber cómo igualar los valores de las diferentes tablas. La cláusula WHERE especifica cómo hacer esto. En la definición de la vista EDITORES_DISCO_COMPACTO, el valor en la columna ID_DISQUERA de la tabla INVENTARIO_CD debe ser igual al valor en la columna ID_DISQUERA de la tabla DISQUERAS para que una fila se devuelva. Por ejemplo, si se refiere de nuevo a la figura 5-3, observará que la fila Past Light en la tabla INVENTARIO_CD tiene un valor de 5412 en la columna ID_DISQUERA, que corresponde a la fila Windham Hill Records en la tabla DISQUERAS. Observe que, una vez más, los nombres de las columnas se califican por los nombres de la tabla, que es esencial en este caso, ya que las columnas comparten el mismo nombre. Sin los nombres de la tabla, SQL no sabría si se comparan valores por sí mismos o con la otra tabla.

También se puede expandir la cláusula WHERE para mayor calidad de consulta. En el siguiente ejemplo, la cláusula WHERE limita las filas devueltas a sólo aquellas que contienen el valor de 5403 en la columna ID_DISQUERA de la tabla INVENTARIO_CD:

```
CREATE VIEW EDITORES_DISCO_COMPACTO
( DISCO_COMPACTO, EDITOR ) AS
  SELECT INVENTARIO_CD.TITULO_CD, DISQUERAS.NOMBRE_COMPANIA
    FROM INVENTARIO_CD, DISQUERAS
   WHERE INVENTARIO_CD.ID_DISQUERA = DISQUERAS.ID_DISQUERA
     AND INVENTARIO_CD.ID_DISQUERA = 5403;
```

Entonces si se invocara la vista EDITORES_DISCO_COMPACTO, se verían sólo los CD que se producen por MCA Records.

Ahora veamos otro ejemplo, que se basa en la vista de la figura 5-4. Del mismo modo que los primeros dos ejemplos que se consideraron, esta vista obtiene los datos de sólo una tabla. Sin embargo, esta vista, de hecho, realiza cálculos que devuelven datos que fueron modificados. La instrucción CREATE VIEW tiene este aspecto:

```
CREATE VIEW DESCUENTOS_CD
( DISCO_COMPACTO, PRECIO_MENUDEO, PRECIO_DESCUENTO ) AS
  SELECT TITULO_CD, PRECIO_MENUDEO, PRECIO_MENUDEO * 0.9
    FROM INVENTARIO;
```

La vista DESCUENTOS_CD incluye tres columnas: DISCO_COMPACTO, PRECIO_MENUDEO y PRECIO_DESCUENTO. La columna PRECIO_DESCUENTO contiene los valores calculados. La cláusula SELECT identifica las columnas de la tabla que contienen los datos fuente. Las dos primeras columnas se definen de la misma manera como se vio en los ejemplos anteriores. Los datos se copian de las columnas TITULO_CD y PRECIO_MENUDEO en la tabla INVENTARIO a las columnas DISCO_COMPACTO y PRECIO_MENUDEO de la vista DESCUENTOS_CD. Sin embargo, la definición de la tercera columna ($\text{PRECIO_MENUDEO} * 0.9$) es un poco diferente. Los valores se toman nuevamente de la columna PRECIO_MENUDEO, sólo que esta vez los valores se multiplican por 0.9 (90 por ciento) para llegar al precio de descuento que aparece en la columna PRECIO_DESCUENTO de la vista.

También se puede añadir la cláusula WHERE a la instrucción SELECT utilizada en la definición de la vista DESCUENTOS_CD:

```
CREATE VIEW DESCUENTOS_CD
( DISCO_COMPACTO, PRECIO_MENUDEO, PRECIO_DESCUENTO ) AS
  SELECT TITULO_CD, PRECIO_MENUDEO, PRECIO_MENUDEO * 0.9
    FROM INVENTARIO
   WHERE EN_EXISTENCIA > 10;
```

La cláusula WHERE restringe la consulta a sólo aquellas filas cuyo valor EN_EXISTENCIA es mayor que 10. Observe que puede utilizar un operador de comparación en una columna de tabla (EN_EXISTENCIA) cuyos valores no son ni siquiera los devueltos por la vista.

Como se puede observar de todos estos ejemplos de definiciones de vista, hay muchas cosas que se pueden hacer con las vistas como resultado de la flexibilidad y extensibilidad de la instrucción SELECT.

Más adelante en el libro, cuando se familiarice más con los distintos tipos de instrucciones `SELECT` que puede crear y las operaciones que puede realizar, será capaz de crear vistas que son mucho más complejas que las que se han considerado hasta ahora.

Creación de vistas actualizables

En SQL, algunos tipos de vistas son actualizables. En otras palabras, se puede utilizar una vista para modificar los datos (cambiar los datos existentes y/o insertar nuevas filas) en la tabla fundamental. Si una vista es actualizable depende de la instrucción `SELECT` que se define dentro de la definición de la vista. Normalmente, entre más compleja es la instrucción, es menos probable que la vista sea actualizable. No hay una sintaxis dentro de la instrucción `CREATE VIEW` que designe explícitamente a una vista como actualizable. En cambio, se determina estrictamente por la naturaleza de la instrucción `SELECT`, que se debe adherir a los estándares específicos para que la vista sea actualizable.

Hasta este punto en el capítulo, se dijo que el marcador de posición <expresión de consulta> en la sintaxis `CREATE VIEW` está compuesto de la instrucción `SELECT`. Para ser más preciso, una expresión de consulta puede ser uno de los varios tipos de expresiones. La más común de éstas, y a la que se refiere este libro, es la especificación de consulta. Una *especificación de consulta* es una expresión de SQL que comienza con la palabra clave `SELECT` e incluye una serie de elementos que forman esa expresión, como se puede observar en los ejemplos de vistas que se examinan. Una especificación de consulta es actualizable si cumple con las numerosas directrices que se exponen en el estándar SQL:2006. En aras de la simplicidad, me refiero a la especificación de consulta como la instrucción `SELECT`, que a menudo es la forma en que se hace referencia en los distintos tipos relacionados con SQL y la documentación relacionada con productos.

Aparte de la cuestión de las especificaciones de consulta y la complejidad de los estándares SQL, el punto que se intenta tratar es que las normas de sintaxis que determinan la actualización de una vista no son simples, directrices claras, en particular teniendo en cuenta el hecho de que aún se tiene que cubrir la instrucción `SELECT` en profundidad (que se hará al comienzo del capítulo 7). Sin embargo, hay bases lógicas que pueden deducirse de estas directrices:

- Los datos sin una vista no pueden ser resumidos, agrupados o eliminados automáticamente.
- Por lo menos una columna de la tabla fuente debe ser actualizable.
- Cada columna en la vista debe ser trazable exactamente a una columna fuente en una tabla.
- Cada fila en la vista debe ser trazable exactamente a una fila fuente en una tabla. Sin embargo, note que muchos proveedores de productos permiten modificaciones (pero no insertan) a las vistas creadas a partir de varias tablas, siempre y cuando la actualización sólo haga referencia a columnas que tracen a sólo una tabla base.

En muchos casos se podrá determinar la actualización de una vista simplemente aplicando sentido común. Veamos un ejemplo. Suponga que decide agregar información acerca de los empleados a la base de datos, ya que quiere rastrear las comisiones ganadas de la venta de CD. Decide crear la tabla `COMISIONES_EMPLEADO`, mostrada en la figura 5-5, que enumera la cantidad total de comisiones por cada empleado hecha durante un periodo de 3 años.

COMISIONES_EMPLEADO

ID_EMPLEADO: INT	AÑO_1999: NUMERICO(7,2)	AÑO_2000: NUMERICO(7,2)	AÑO_2001: NUMERICO(7,2)
99301	126.32	11.28	16.86
99302	16.27	90.20	198.14
99303	354.34	16.32	1237.56
99304	112.11	87.56	14.14

Figura 5-5 Ingresos anuales por comisión en la tabla COMISIONES_EMPLEADO.

Ahora suponga que quiere saber el promedio de la comisión por cada año para todos los empleados. Se puede crear una vista que determine el promedio por cada año y muestre esos promedios en tres columnas separadas. Para hacerlo se utiliza la siguiente instrucción CREATE VIEW:

```
CREATE VIEW COM_EMP ( PROM_1999, PROM_2000, PROM_2001 ) AS
  SELECT PROM(AÑO_1999), PROM(AÑO_2000), PROM(AÑO_2001)
  FROM COMISIONES_EMPLEADO;
```

Como se puede ver de la instrucción, la vista COM_EMP contiene tres columnas: PROM_1999, PROM_2000 y PROM_2001. La cláusula SELECT extrae la información de las tres columnas en la tabla COMISIONES_EMPLEADO (AÑO_1999, AÑO_2000 y AÑO_2001) y utiliza la función PROM para encontrar el promedio para todos los valores en cada columna, como se muestra en la figura 5-6. Por ejemplo, la función PROM primero promedia los cuatro valores en la columna AÑO_1999 y luego introduce ese promedio en la columna PROM_1999 de la vista COM_EMP.

Ahora suponga que desea actualizar los importes de la comisión en la tabla COMISIONES_EMPLEADO. No se podría hacer a través de la vista, ya que los valores en la vista se basan en cálculos realizados sobre los valores en la tabla. Por ejemplo, si se trata de actualizar los valores en la columna PROM_1999, los RDBMS no sabrían cómo se afectan muchas de las filas o cómo se distribuyen los valores dentro de esas filas. En otras palabras, la fila en la vista no se puede rastrear de nuevo a exactamente la fila fuente.

COM_EMP

PROM_1999	PROM_2000	PROM_2001
152.26	51.34	366.68

Figura 5-6 La vista COM_EMP, basada en el promedio trimestral de ingresos.

COM_EMP		
ID_EMPLEADO	AÑO_1999	AÑO_2000
99301	126.32	11.28
99302	16.27	90.20
99303	354.34	16.32
99304	112.11	87.56

Figura 5-7 Vista COM_EMP, basada en el primero y segundo trimestres de ingresos.

Se puede, sin embargo, crear una vista que simplemente extraiga información de la tabla COMISIONES_EMPLEADO:

```
CREATE VIEW COM_EMP AS
SELECT ID_EMPLEADO, AÑO_1999, AÑO_2000
FROM COMISIONES_EMPLEADO;
```

En esta instrucción se crea una vista que muestra sólo tres de las cuatro columnas de la tabla. No se realizan cálculos y sólo se utiliza una tabla. La figura 5-7 muestra cómo se ve la vista.

Esta vista, a diferencia de la última, es actualizable. Se pueden modificar e insertar datos, ya que éstos no se resumieron o agruparon; cada columna se puede rastrear exactamente a una columna fuente en una tabla; cada fila se rastrea exactamente a una fila fuente en una tabla. Además, ningún dato se resume o se agrupa. Por supuesto, si se fueran a actualizar o insertar datos mediante la vista, son los datos en la tabla fundamental los que en realidad se modifican. Eso significa que cualquier modificación de datos debe aún adherirse a las limitaciones que tienen lugar en la tabla. Por ejemplo, no se podría insertar una fila a través de la vista COM_EMP si los valores nulos no fueran permitidos en la columna AÑO_2001 de la tabla. La vista no tendría la capacidad de aceptar un valor para esa columna, y la tabla no permitiría insertar una fila sin suministrar ese valor.

A menudo se puede determinar si una tabla es actualizable sólo observando el resultado de cualquier intento de modificación. Si su objetivo es crear vistas que permitan a los usuarios actualizar datos en las tablas fundamentales, entonces debe considerar la complejidad de esas vistas y las funciones que vayan a realizar. También considere que las limitaciones que tengan lugar en las tablas fundamentales afectan la habilidad para modificar e insertar datos mediante una vista.

Utilice la cláusula WITH CHECK OPTION

Ahora volvamos a la sintaxis de CREATE VIEW que ya se presentó en la sección “Definición de las vistas de SQL”. La última línea de la sintaxis incluye la siguiente cláusula:

[WITH CHECK OPTION]

La cláusula `WITH CHECK OPTION` aplica a vistas actualizables que incluyen la cláusula `WHERE` en la instrucción `SELECT`. La mejor manera de ilustrar cómo funciona esto es por medio de un ejemplo. Modifiquemos la definición de la última vista que examinamos:

```
CREATE VIEW COM_EMP AS
  SELECT ID_EMPLEADO, AÑO_1999, AÑO_2000
  FROM COMISIONES_EMPLEADO
  WHERE AÑO_1999 > 100;
```

La cláusula `WHERE` especifica que sólo las filas `AÑO_1999` con valores mayores que 100 se devuelven. Esto en sí es bastante claro. Sin embargo, suponga que quiere actualizar esta vista estableciendo un valor `AÑO_1999` que sea menor o igual que 100. Ya que la vista es actualizable, le permitirá hacer eso. Sin embargo, si se invoca la vista, la fila que se actualizó ya no sería visible ni se podría actualizar más.

Para evitar este problema se puede agregar la cláusula `WITH CHECK OPTION` a la definición de la vista, como en el siguiente ejemplo:

```
CREATE VIEW COM_EMP AS
  SELECT ID_EMPLEADO, AÑO_1999, AÑO_2000
  FROM COMISIONES_EMPLEADO
  WHERE AÑO_1999 > 100
  WITH CHECK OPTION;
```

Ahora si se trata de actualizar un valor `AÑO_1999` a una cantidad menor o igual que 100, se recibe un mensaje de error diciendo que el cambio no se puede hacer. Como se observa, la cláusula `WITH CHECK OPTION` es una manera práctica de garantizar que los usuarios no realicen actualizaciones que les impida la utilización eficaz de las vistas que se crean.

Eliminación de vistas de la base de datos

No cabe duda que se enfrentará a situaciones en las que desee quitar la definición de una vista de su base de datos. La sintaxis para hacerlo es bastante simple:

```
DROP VIEW <nombre de la vista>
```

Cuando se ejecuta la instrucción `DROP VIEW`, la definición de la vista se quita; sin embargo, ninguno de los datos fundamentales (que se almacenan en las tablas base) se afectan. Una vez que la vista se elimina, se puede volver a crear la vista o crear una vista diferente con el mismo nombre. Ahora veamos un ejemplo rápido:

```
DROP VIEW COM_EMP;
```

Esta instrucción elimina la vista `COM_EMP` de la base de datos, pero deja los datos fundamentales intactos.

Pregunta al experto

P: Se analizó la creación y eliminación de vistas, pero no se mencionó su modificación. ¿Respalda SQL cualquier tipo de instrucción ALTER VIEW?

R: No, el estándar SQL:2006 no soporta la modificación de las vistas. Sin embargo, algunos RDBMS soportan la instrucción ALTER VIEW. Sin embargo, esté consciente de que la funcionalidad respaldada por estas instrucciones puede variar de un producto a otro. Por ejemplo, SQL Server y MySQL tienen las instrucciones ALTER VIEW, que son bastante robustas y permiten cambiar muchos aspectos de la definición de una vista, incluyendo la instrucción SELECT. Por otra parte, la instrucción ALTER VIEW en Oracle se utiliza para recompilar manualmente, a fin de evitar la sobrecarga en el tiempo de ejecución o modificar algunas limitaciones que Oracle respalda en las vistas. De hecho, para modificar una vista en Oracle tiene primero que eliminarse y después volver a crearse, como es el caso en el estándar SQL. Sin embargo, Oracle tiene la instrucción CREATE OR REPLACE VIEW que esencialmente permite eliminar y volver a crear una vista en un solo paso.

P: En los ejemplos que se usaron para mostrar cómo se crean las vistas, se utilizaron una o dos tablas para los datos fuente. ¿Se pueden basar las vistas en más de dos tablas?

R: Sí, una vista puede basarse en tantas tablas como pueda ser lógicamente consultada en la instrucción SELECT. Por ejemplo, suponga que desea crear una vista en la base de datos INVENTARIO. (La base de datos INVENTARIO es con la que se ha trabajado para los ejercicios en el libro.) La vista muestra los nombres de los artistas junto con los títulos de CD. Sin embargo, para hacer esto, la instrucción SELECT tiene que unir tres tablas. Tendrían que coincidir los valores de ID_ARTISTA en la tabla ARTISTAS y la tabla CD_ARTISTA, y tendrían que coincidir los valores ID_DISCO_COMPACTO en la tabla DISCOS_COMPACTOS y la tabla CD_ARTISTA. El resultado sería una vista que muestra una lista de artistas y sus CD. (En el capítulo 11 se analizará cómo se pueden unir estas tablas en la instrucción SELECT.)

P: En los ejemplos que se utilizaron para mostrar cómo se crean las vistas, todas las vistas hacen referencia a tablas base. ¿Se crean todas las vistas a partir de tablas base?

R: No, las vistas se pueden crear utilizando expresiones de consulta que extraen datos de otras vistas. También es posible crear una vista que contenga sólo datos calculados y, por lo tanto, no disponga de datos que dirijan a una tabla base.

Pruebe esto 5-1 **Añada vistas a su base de datos**

En este ejercicio se crearán dos vistas en la base de datos INVENTARIO. Las vistas se basan en las tablas que se crearon en los ejercicios anteriores. La primera vista se basa en una sola tabla, y la segunda vista en dos tablas. La segunda vista se creará dos veces. Se creará una vez, luego se eliminará la definición de la vista de la base de datos, y a continuación se volverá a crear una versión modificada de la vista. Se puede descargar el archivo Try_This_05.txt, que contiene las instrucciones SQL utilizadas en este ejercicio (en inglés).

Paso a paso

1. Abra la aplicación de cliente de su RDBMS y conéctese a la base de datos INVENTARIO.
2. La primera vista que se crea se llama CD_EN_EXISTENCIA. La vista se basa en las columnas TITULO_CD y EN_EXISTENCIA en la tabla DISCOS_COMPACTOS. Se desea que la vista incluya sólo las filas cuyos valores en la columna EN_EXISTENCIA sean mayores que 10. La vista utilizará los mismos nombres de columna que la tabla e incluirá la cláusula WITH CHECK OPTION para evitar que los valores inferiores o iguales a 10 se añadan a la columna EN_EXISTENCIA. Introduzca y ejecute la siguiente instrucción SQL:

```
CREATE VIEW CDS_EN_EXISTENCIA AS
  SELECT TITULO_CD, EN_EXISTENCIA
    FROM DISCOS_COMPACTOS
   WHERE EN_EXISTENCIA > 10 WITH CHECK OPTION;
```

3. Después se crea la vista llamada EDITORES_CD, que contiene las columnas TITULO_CD y EDITOR. La vista se basa en la columna TITULO_CD de la tabla DISCOS_COMPACTOS y en la columna NOMBRE_COMPañIA de la tabla DISQUERAS_CD. Tendrá que utilizar la cláusula WHERE para que coincidan las filas en las dos tablas. La cláusula WHERE también va a limitar las filas incluidas en la vista a aquellas cuyo valor ID_DISQUERA en la tabla DISQUERAS_CD sea 5403 o 5402. Introduzca y ejecute la siguiente instrucción SQL:

```
CREATE VIEW EDITORES_CD
  ( TITULO_CD, EDITOR ) AS
  SELECT DISCOS_COMPACTOS.TITULO_CD, DISQUERAS_CD.NOMBRE_COMPañIA
    FROM DISCOS_COMPACTOS, DISQUERAS_CD
   WHERE DISCOS_COMPACTOS.ID_DISQUERA = DISQUERAS_CD.ID_DISQUERA
     AND DISQUERAS_CD.ID_DISQUERA = 5403 OR DISQUERAS_CD.ID_DISQUERA =
5402;
```

4. Decide que no desea limitar las filas a valores específicos en la columna ID_DISQUERA, así que debe eliminar la definición de la vista de la base de datos y volver a crear la vista sin las restricciones en los valores. Introduzca y ejecute la siguiente instrucción SQL:

```
DROP VIEW EDITORES_CD;
```

(continúa)

5. Ahora se puede volver a crear la vista EDITORES_CD. Introduzca y ejecute la siguiente instrucción SQL:

```
CREATE VIEW EDITORES_CD
( TITULO_CD, EDITOR ) AS
SELECT DISCOS_COMPACTOS.TITULO_CD, DISQUERAS_CD.NOMBRE_COMPANÍA
FROM DISCOS_COMPACTOS, DISQUERAS_CD
WHERE DISCOS_COMPACTOS.ID_DISQUERA = DISQUERAS_CD.ID_DISQUERA;
```

6. Cierre la aplicación de cliente.

Resumen de Pruebe esto

Además de las seis tablas creadas en los ejercicios anteriores, la base de datos ahora incluye las vistas CD_EN_EXISTENCIA y EDITORES_CD. Más adelante en el libro, se utilizarán estas vistas para consultar datos de las tablas base y actualizar esos datos. Una vez que tenga una mejor comprensión de cómo crear las instrucciones SELECT, será capaz de definir vistas que son aún más amplias y proporcionan un mayor nivel de detalle que las vistas que se crearon hasta ahora.

Autoexamen Capítulo 5

1. ¿Cuáles son las dos ventajas de utilizar vistas?
2. ¿Cuáles son los tres tipos de tablas almacenadas, soportadas por SQL?
3. ¿Qué sucede si no se asignan nombres de columna a una vista?
4. ¿Cómo se asignan los tipos de datos a las columnas de una vista?
5. ¿En qué circunstancias se deben proporcionar nombres a las columnas de una vista en una definición de vista?
6. Se crea la vista llamada EMP_CUMPLEAÑOS. La vista se basa en las columnas EMP_NOMBRE y CUMPLEAÑOS de la tabla EMPLEADOS. Los nombres de columna de la vista serán los mismos que los nombres de columna de la tabla. ¿Qué código SQL se utilizaría para crear la vista?
7. Se crea una vista basada en la tabla DISCOS_COMPACTOS en la base de datos INVENTARIO. Se desea que la vista incluya sólo las filas cuyo valor en la columna ID_DISQUERA sea 546. ¿Qué cláusula (además de las cláusulas SELECT y FROM) deberá incluirse en la instrucción SELECT para la vista?
8. Se crea una vista que hace referencia a las tablas EMPLEADOS y TITULO_TRABAJO. Los datos en las dos tablas coinciden a través de la columna ID_TITULO_TRABAJO en cada tabla. ¿Cómo se debe escribir la cláusula WHERE en la instrucción SELECT de la vista?

9. Se crea una vista que hace referencia a las tablas EMPLEADOS y TITULO_TRABAJO. Los datos en las dos tablas coinciden por la columna ID_TITULO_TRABAJO en cada tabla. Se desea que la vista muestre sólo las filas cuyo valor en la columna ID_TITULO_TRABAJO de la tabla TITULO_TRABAJO sea 109. ¿Cómo se debe escribir la cláusula WHERE en la instrucción SELECT de la vista?
10. ¿Qué es una especificación de consulta?
11. ¿Qué directrices deben seguirse si desea crear una vista actualizable?
 - A Los datos dentro de la vista no se pueden resumir, agrupar o eliminar automáticamente.
 - B Por lo menos una columna en la tabla fuente debe ser actualizable.
 - C Cada columna en la vista debe rastrearse exactamente a una columna fuente en una tabla.
 - D Cada fila en la vista debe rastrearse exactamente a una fila fuente en una tabla.
12. Se crea la siguiente vista basada en la tabla DISCOS_COMPACTOS en la base de datos INVENTARIO:

```
CREATE VIEW PROMEDIO_EN_EXISTENCIA AS
SELECT PROM(EN_EXISTENCIA)
FROM DISCOS_COMPACTOS;
```

¿Cómo se insertan datos a través de esta vista?

13. ¿A qué tipo de vista aplica la cláusula WITH CHECK OPTION?
14. Se crea la siguiente definición de vista:

```
CREATE VIEW COM_EMP AS
SELECT ID_EMPLEADO, AÑO_1999, AÑO_2000
FROM COMISION_EMPLEADO
WHERE AÑO_1999 > 100;
```

Se quiere utilizar la vista para actualizar los datos. ¿Qué sucede si se cambia el valor AÑO_1999 a una cantidad inferior o igual que 100?

15. Se desea modificar la definición de la vista COM_EMP en la base de datos. ¿Cómo se modifica esa definición?
16. Se desea eliminar la definición de la vista EMP_CUMPLEAÑOS de la base de datos. ¿Qué instrucción SQL deberá utilizarse?
17. ¿Qué les sucede a los datos de SQL cuando se elimina una vista de la base de datos?

Capítulo 6

Gestión de seguridad en la base de datos

Habilidades y conceptos clave

- Entienda el modelo de seguridad de SQL
 - Creación y eliminación de roles
 - Otorgue y revoke privilegios
 - Otorgue y revoke roles
-

Un componente crítico de cualquier base de datos es la habilidad para proteger los datos contra accesos no autorizados y ataques maliciosos. Una base de datos debe garantizar que los usuarios no autorizados no puedan ver o cambiar datos que no deban ver o cambiar. Al mismo tiempo, a los usuarios autorizados no se les debe impedir acceder a cualquier información que está disponible para ellos. El equilibrio ideal es dar con exactitud a cada uno de los usuarios de base de datos los privilegios que necesitan para hacer su trabajo, nada más y nada menos. Con el fin de respaldar estas capacidades, SQL define un modelo de seguridad que permite determinar qué usuarios pueden acceder a datos específicos y lo que pueden hacer con esos datos. La parte esencial de este modelo es el identificador de autorización. Un *identificador de autorización*, como se aprendió en el capítulo 2, es un objeto en el entorno SQL que representa a un usuario o grupo de usuarios a los que se les otorgan privilegios específicos para acceder a objetos y datos dentro del entorno SQL. Los privilegios a los objetos de esquema se les conceden a los identificadores de autorización. El tipo de privilegio otorgado determina el tipo de acceso. En este capítulo se examina el modelo de seguridad de SQL, cómo utiliza a los identificadores de autorización y cómo configura los privilegios sobre objetos en la base de datos SQL.

Entienda el modelo de seguridad de SQL

Los identificadores de autorización proporcionan la base para la seguridad en la base de datos. Se permite acceder a todos los objetos a través de estos identificadores. Si el identificador de autorización no tiene los privilegios apropiados para acceder a un objeto específico, como una tabla, los datos dentro de esa tabla no están disponibles para ese usuario. Además, cada identificador de autorización se puede configurar con diferentes tipos de privilegios. Por ejemplo, puede permitir que algunos identificadores de autorización consulten los datos dentro de una tabla específica, mientras que permita a otros identificadores de autorización tanto ver como modificar esos datos.

SQL respalda dos tipos de identificadores de autorización: identificadores de usuario (o usuarios) y nombres de rol (o roles). Un *identificador de usuario* es una cuenta de seguridad individual que puede representar a una persona, una aplicación o un servicio del sistema (de los cuales todos se consideran como usuarios de la base de datos). El estándar SQL no especifica cómo una aplicación de SQL puede crear a un identificador de usuario. El identificador puede estar vinculado al sistema operativo en el que se ejecuta el sistema de gestión de base de datos relacional (RDBMS), o puede estar creado explícitamente en el entorno RDBMS.

Un *nombre de rol* es un conjunto de privilegios definidos que se pueden asignar a un usuario o a otro rol. Si a un nombre de rol se le concede acceder a un objeto de esquema, entonces todos los identificadores de usuario y los nombres de rol que se asignaron al nombre del rol específico se les concede acceder a ese objeto siempre y cuando el nombre de rol sea el del identificador de autorización actual. Por ejemplo, en la figura 6-1 el nombre de rol DEPTO_MRKT se asignó al nombre de rol DEPTO_ACCT y los cuatro identificadores de usuario: Ethan, Max, Linda y Emma. Si el nombre de rol DEPTO_ACCT es el identificador de autorización actual y se le concedió acceder a la tabla INTERPRETES, el nombre de rol DEPTO_ACCT y los cuatro identificadores de usuario tienen acceso a la tabla INTERPRETES. Observe que, a diferencia de un identificador de usuario, SQL *sí* especifica cómo crear un nombre de rol, que se analizará en la sección “Creación y eliminación de roles”, más adelante en este capítulo.

Los nombres de rol se utilizan comúnmente como un mecanismo para la concesión de un conjunto uniforme de privilegios a los identificadores de autorización que deben tener los mismos privilegios, como las personas que trabajan en el mismo departamento. También tienen la ventaja de la existencia independiente de los identificadores de usuario, lo que significa que se pueden crear antes que los identificadores de usuario, y persisten incluso después de que los identificadores de usuario suprimen las referencias. Esto es muy útil a la hora de administrar los privilegios para un trabajo fluido.

Además de los identificadores de usuario y los nombres de rol, SQL respalda a un identificador de autorización especial incorporado llamado PUBLIC, que incluye a todos los que utilizan la base de datos. Al igual que con cualquier otro identificador de autorización, se pueden conceder privilegios de acceso a la cuenta PUBLIC. Por ejemplo, suponga que quiere que todos los clientes potenciales puedan ver su lista de CD.

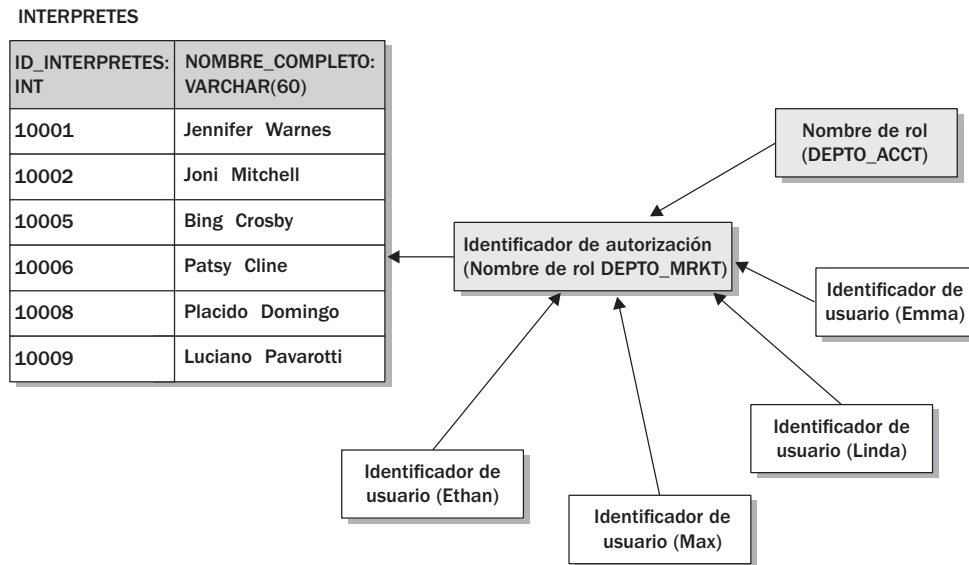


Figura 6-1 Rol DEPTO_MRKT asignado a cuatro identificadores de usuario y a un rol.

Puede conceder los privilegios necesarios a la cuenta PUBLIC para las tablas y columnas correspondientes. Evidentemente, la cuenta PUBLIC debe utilizarse con sumo cuidado, ya que puede abrir la oportunidad para personas con malas intenciones. De hecho, muchas organizaciones prohibieron su uso por completo.

Sesiones SQL

Cada sesión SQL se asocia con un identificador de usuario y un nombre de rol. Una *sesión SQL* es la conexión entre algún tipo de aplicación de cliente y la base de datos. La sesión proporciona el contexto en el que el identificador de autorización ejecuta las instrucciones SQL durante una sola conexión. A través de esta conexión, la sesión SQL mantiene la asociación con el par identificador de usuario / nombre de rol.

Echemos un vistazo a la figura 6-2, que muestra al par identificador de usuario/nombre de rol asociado con una sesión. Cuando una sesión se establece por primera vez, el identificador de usuario es siempre el *identificador de usuario de la sesión SQL*, que es un tipo especial de identificador de usuario que permanece asociado con la sesión a través de la conexión. Corresponde a la aplicación SQL determinar cómo una cuenta específica se convierte en el identificador de usuario de la sesión SQL, aunque puede ser una cuenta de usuario del sistema operativo o una cuenta específica para el RDBMS. Cualquier método que se utilice para asociar una cuenta con el identificador de usuario de la sesión SQL, esta cuenta es la que actúa como identificador de usuario actual.

Como también se puede observar en la figura 6-2, el nombre de rol es un valor nulo. El nombre de rol es siempre nulo cuando una sesión se establece por primera vez. En otras palabras, cada vez que la base de datos de SQL inicie y establezca una sesión, el identificador de usuario inicial siempre será el identificador de usuario de la sesión SQL y el nombre de rol siempre será un valor nulo.

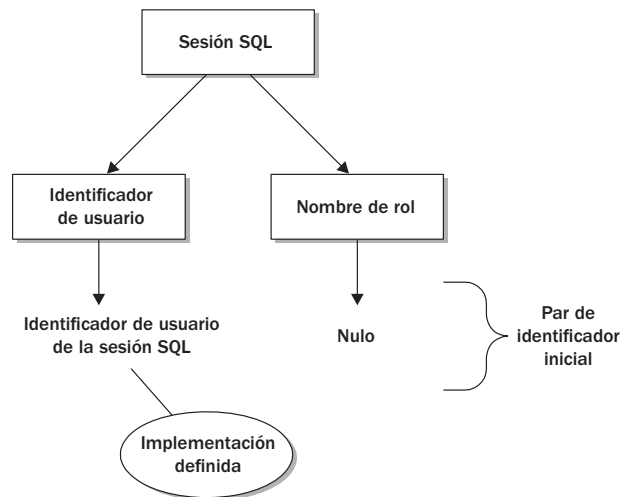


Figura 6-2 Sesión SQL con un identificador de usuario y un nombre de rol.

En cualquier momento durante una conexión, la sesión se asocia con un par identificador de usuario/nombre de rol; sin embargo, no es siempre el mismo par a lo largo del periodo de la sesión. Por ejemplo, las instrucciones SQL incrustadas, los módulos de clientes SQL y las rutinas invocadas SQL pueden especificar un identificador de autorización. Si se especifica un identificador nuevo, se convierte en el identificador de autorización actual hasta que la transacción se complete, y el acceso a los objetos se concede sobre la base del par actual identificador de usuario/nombre de rol.

Para cualquier par actual de identificador de usuario/nombre de rol, uno de los dos valores casi siempre es nulo. En otras palabras, si se especifica un identificador de autorización, entonces el nombre de rol debe ser nulo; si un nombre de rol se especifica, entonces el identificador de usuario debe ser nulo. Cualquier valor que no sea nulo es el identificador de autorización.

Cuando más de un par identificador de usuario/nombre del rol se utiliza durante una sesión, se crea una pila de autorización que refleja al identificador de usuario actual. El par en la parte superior de la pila es el identificador de autorización actual. La figura 6-3 muestra un ejemplo de una pila de autorización que se puede crear durante una sesión.

En este ejemplo, el par identificador de usuario/nombre del rol inicial está en la parte inferior de la pila. Como cabría esperar, el identificador de usuario es el identificador de usuario de la sesión SQL y el nombre de rol es un valor nulo. Acceder a un objeto de base de datos se basa en los privilegios concedidos al identificador de usuario de la sesión SQL cuando es actual.

Durante la sesión, una instrucción SQL incrustada especifica un identificador de autorización Usuario_App, que es un identificador de usuario. Cuando la instrucción incrustada se ejecuta, el Usuario_App se convierte en el identificador de autorización actual y los privilegios de acceso se basan en esa cuenta.

Suponga una de las instrucciones SQL incrustadas, luego llame una rutina SQL invocada que especifique una autorización de ROL_RUTINA, que es un nombre de rol. ROL_RUTINA se convierte en el identificador de autorización actual y está en la parte superior de la pila de autorización. Una vez que se ejecuta la ruta, el identificador de autorización actual revierte al Usua-

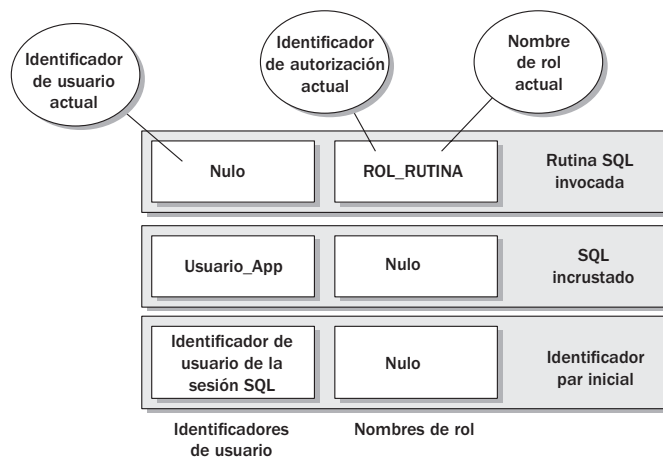


Figura 6-3 Pila de autorización creada durante una sesión SQL.

rio_App, hasta que las instrucciones incrustadas se ejecuten, con lo que después el identificador de autorización se revierte al identificador de usuario de la sesión SQL.

Observe que en cada par identificador de usuario/nombre de rol mostrado en la figura 6-3 hay exactamente un valor nulo. El otro valor, el que no es nulo, es el identificador de autorización.

Pregunta al experto

- P:** Se expuso que el identificador de autorización actual puede cambiar. ¿Cómo se puede determinar la autorización del usuario actual y el nombre de rol en cualquier momento durante una sesión?
- R:** SQL soporta varios valores especiales que permiten determinar los valores actuales de varios tipos de usuarios. Los valores especiales actúan como marcadores de posición para el valor de usuario relacionado actual. Se pueden utilizar estos valores especiales en expresiones para devolver el valor del tipo específico de usuario. Por ejemplo, se puede utilizar el valor especial `USUARIO_ACTUAL` para devolver el valor del identificador de usuario actual. SQL respalda cinco de estos valores especiales: `USUARIO_ACTUAL`, `USUARIO`, `ROL_ACTUAL`, `USUARIO_SESION` y `USUARIO_SISTEMA`. Los valores `USUARIO_ACTUAL` y `USUARIO` significan lo mismo y devuelven un valor igual al del identificador de usuario actual. El `ROL_ACTUAL` devuelve el nombre de rol actual, y `USUARIO_SESION` devuelve el identificador de usuario de la sesión SQL. Si el identificador de usuario de la sesión SQL es el identificador de usuario actual, entonces `USUARIO_ACTUAL`, `USUARIO` y `USUARIO_SESION` tienen el mismo valor, que puede ocurrir si el par identificador inicial es el único par identificador de usuario/nombre de rol activo (el par en la parte superior de la pila de autorización). La última función, `USUARIO_SISTEMA`, devuelve el sistema operativo del usuario que invocó el módulo SQL. A medida que avancemos más en este capítulo, se verá cómo los valores especiales `USUARIO_ACTUAL` y `ROL_ACTUAL` se utilizan para reconocer al identificador de autenticación actual en el momento de crear los roles y conceder privilegios. (Ver las secciones “Creación y eliminación de roles”, “Otorgue y revoque privilegios” y “Otorgue y revoque roles”.) Además, se encuentra más información sobre valores especiales en el capítulo 10.

Acceda a objetos de base de datos

Ahora que tiene una mejor comprensión de lo que es un identificador de autorización (junto con los identificadores de usuario y los nombres de rol), echemos un vistazo de lo que puede hacer con estos identificadores. Acceder a los datos en una base de datos se basa en la posibilidad de acceder a los objetos que contienen los datos. Por ejemplo, puede conceder a algunos usuarios el acceso a un conjunto específico de tablas, mientras que otros usuarios pueden acceder sólo a columnas específicas dentro de una tabla. SQL permite definir los privilegios de acceso a los siguientes objetos de esquema:

- Tablas base
- Vistas

- Columnas
- Dominios
- Conjunto de caracteres
- Cotejos
- Traducciones
- Tipos de usuarios definidos
- Secuencias
- Activadores
- Rutas invocadas SQL

Para cada tipo de objeto se pueden asignar determinados tipos de privilegios que varían según el tipo de objeto. Estos privilegios asignados se asocian con identificadores de autorización específicos. En otras palabras, se pueden asignar uno o más de los privilegios de un objeto a uno o más identificadores de autorización. Por ejemplo, se puede asignar el privilegio SELECT para una tabla para el identificador de autorización PUBLIC. Esto permite a todos los usuarios de la base de datos ver los contenidos de esa tabla.

SQL define nueve tipos de privilegios que se pueden asignar a un objeto de esquema. La tabla 6-1 describe cada uno de estos privilegios y enumera los tipos de objetos a los cuales se puede asignar el privilegio.

Los privilegios se conceden en las bases de datos objetos utilizando la instrucción GRANT para especificar los objetos así como los identificadores de autorización que adquieren los privile-

Privilegio	Descripción	Objetos
SELECT	Permite que identificadores de autorización específicos consulten datos en el objeto. Por ejemplo, si al UsuarioA se le concede el privilegio SELECT en la tabla CD_ARTISTAS, el usuario puede ver los datos de esa tabla.	Tablas Vistas Columnas Métodos (en tipos estructurados)
INSERT	Permite que identificadores de autorización específicos inserten datos en el objeto. Por ejemplo, si al UsuarioA se le concede el privilegio INSERT en la tabla CD_ARTISTAS, el usuario puede añadir datos a esa tabla.	Tablas Vistas Columnas
UPDATE	Permite que identificadores de autorización específicos actualicen datos en el objeto. Por ejemplo, si al UsuarioA se le concede el privilegio UPDATE en la tabla CD_ARTISTAS, el usuario puede modificar datos a esa tabla. Sin embargo, este privilegio no le permite al usuario cambiar la definición de la tabla.	Tablas Vistas Columnas
DELETE	Permite que identificadores de autorización específicos eliminen datos del objeto. Por ejemplo, si al UsuarioA se le concede el privilegio DELETE en la tabla CD_ARTISTAS, el usuario puede eliminar datos de esa tabla. Sin embargo, este privilegio no le permite al usuario eliminar la definición de la tabla de la base de datos.	Tablas Vistas

Tabla 6-1 Privilegios de seguridad asignados a objetos de base de datos.

Privilegio	Descripción	Objetos
REFERENCES	Permite que identificadores de autorización específicos definan los objetos (como limitaciones referenciales) que hacen referencia a la tabla configurada con el privilegio REFERENCES. Por ejemplo, si al UsuarioA se le concede el privilegio REFERENCES en la tabla CD_ARTISTAS, el usuario puede crear otros objetos que hagan referencia a la tabla CD_ARTISTAS, como sería el caso con claves foráneas. (Note que el UsuarioA también debe tener la autorización para crear otros objetos.)	Tablas Vistas Columnas
TRIGGER	Permite que identificadores de autorización específicos generen activadores en la tabla. Por ejemplo, si al UsuarioA se le concede el privilegio TRIGGER en la tabla CD_ARTISTAS, el usuario puede crear activadores en esa tabla.	Tablas
USAGE	Permite que los identificadores de autorización específicos utilicen el objeto en una definición de columna. Por ejemplo, si al UsuarioA se le concede el privilegio USAGE en el dominio DINERO, el usuario puede incluir el dominio en la definición de columna cuando se crea una tabla. (Note que el UsuarioA también debe tener la autorización para crear una tabla.)	Dominios Conjunto de caracteres Cotejos Traducciones Tipos definidos por el usuario Secuencias
EXECUTE	Permite que los identificadores de autorización específicos invoquen una rutina SQL invocada. Por ejemplo, si al UsuarioA se le concede el privilegio EXECUTE en el procedimiento almacenado LISTA_CD_ACTUALIZADA, el usuario sería capaz de invocar ese procedimiento almacenado.	Rutinas invocadas SQL
UNDER	Permite que los identificadores de autorización específicos definan un subtipo directo en un tipo estructurado. Un <i>subtipo directo</i> es un tipo estructurado que se asocia con otro tipo estructurado como un objeto secundario de ese tipo. Por ejemplo, si al UsuarioA se le concede el privilegio UNDER en el tipo estructurado EMPLEADO, el usuario puede definir subtipos directos tales como ADMINISTRADOR o SUPERVISOR.	Tipos estructurados

Tabla 6-1 Privilegios de seguridad asignados a objetos de base de datos (continuación).

gios. También se pueden revocar privilegios usando la instrucción REVOKE. Se entrará en mayor detalle sobre ambas instrucciones a medida que avancemos por el capítulo. Sin embargo, antes de discutir cómo conceder o revocar privilegios, primero deseo discutir cómo crear un nombre de rol. (Recuerde, el estándar SQL no respalda la creación de un identificador de usuario, sólo nombres de rol. El proceso de creación de identificadores de usuario es la aplicación específica.)

Creación y eliminación de roles

En su mayor parte, la creación de un rol es un proceso muy sencillo. La instrucción incluye sólo una cláusula obligatoria y una cláusula opcional, como se muestra en la siguiente sintaxis:

```
CREATE ROLE <nombre del rol>
[ WITH ADMIN { CURRENT_USER | CURRENT_ROLE } ]
```

Observe que la única parte requerida de la sintaxis es la cláusula `CREATE ROLE`, lo que significa que todo lo que realmente se necesita hacer es especificar un nombre para el rol. La cláusula `WITH ADMIN` es opcional y rara vez necesitará utilizarla. Es necesario sólo si el par identificador de usuario/nombre de rol actual contiene valores no nulos. La cláusula permite diseñar tanto al identificador de usuario actual (`CURRENT_USER`) o el nombre de rol actual (`CURRENT_ROLE`) como el identificador de autenticación permite asignar el rol a los identificadores de usuario o los nombres de rol. Si la cláusula `WITH ADMIN` no se especifica, al identificador de autenticación actual, ya sea el identificador de usuario actual o el nombre de rol actual, se le permite asignar el rol.

NOTA

Probablemente encontrará que rara vez necesite utilizar la cláusula `WITH ADMIN` de la instrucción `CREATE ROLE`, particularmente como programador principiante de SQL, y no es ampliamente respaldado en los productos RDBMS. Como resultado, se mantiene una breve discusión de la cláusula.

Ahora veamos la creación de un rol. En el siguiente ejemplo se utiliza la instrucción `CREATE ROLE` para crear el rol `CLIENTES`:

```
CREATE ROLE CLIENTES;
```

Eso es todo lo que hay que hacer. Una vez que se crea el rol, se puede otorgar el rol a identificadores de usuario u otros nombres de roles. Se analizará la concesión y revocación de roles en la sección “Otorgue y revoque roles” después en este capítulo.

La eliminación de un rol es tan fácil como la creación de uno. La sintaxis que se utiliza es la siguiente:

```
DROP ROLE <nombre del rol>
```

En este caso, simplemente necesita identificar el nombre del rol, como en el siguiente ejemplo:

```
DROP ROLE CLIENTES;
```

El rol se elimina de la base de datos. Sin embargo, antes de eliminar un rol, asegúrese que se trata del rol que ya no es necesario o que es el que particularmente quiere eliminar (por razones de seguridad).

Como se puede observar, la creación y eliminación de roles es un proceso muy sencillo, y puede hacer la gestión de los usuarios mucho más fácil. Los roles esencialmente permiten agrupar los usuarios que requieren los mismos privilegios sobre los mismos objetos. Ahora echemos un vistazo a la concesión y revocación de privilegios para los identificadores de autenticación, incluyendo los identificadores de usuario y los nombres de rol.

NOTA

El respaldo para las instrucciones `CREATE ROLE` y `DROP ROLE` varía de una aplicación a otra. Por ejemplo, Oracle y SQL Server (2005 y 2008) respaldan ambas instrucciones, pero no la opción `WITH ADMIN`. MySQL 5.0 no parece respaldar el concepto de roles.

Otorgue y revoque privilegios

Cuando se conceden privilegios en un objeto, se asocian uno o más privilegios con uno o más identificadores de autorización. Este conjunto de privilegios e identificadores de autorización se

asignan al objeto, que permite al identificador de autorización poder acceder al objeto según el tipo de privilegios definidos. Para otorgar privilegios, se debe utilizar la instrucción `GRANT`, como se muestra en la siguiente sintaxis:

```
GRANT { ALL PRIVILEGES | <lista de privilegios> }  
ON <tipo de objeto> <nombre del objeto>  
TO { PUBLIC | <lista de identificador de autorización> } [WITH GRANT OPTION]  
[GRANTED BY { CURRENT_USER | CURRENT_ROLE }]
```

La instrucción, como se puede observar, incluye tres cláusulas requeridas —`GRANT`, `ON` y `TO`— y dos cláusulas opcionales —`WITH GRANT OPTION` y `GRANTED BY`—. Se analizará cada cláusula individualmente, excepto la cláusula `GRANTED BY`. La cláusula `GRANTED BY` es similar a la cláusula `WITH ADMIN` en la instrucción `CREATE ROLE`. Como esa cláusula, la cláusula `GRANTED BY` aplica sólo en esas situaciones donde el par identificador de usuario/nombre de rol actual contiene valores no nulos, y no se implementa ampliamente en los RDBMS. Como principiante en la programación de SQL, no necesita inquietarse con la cláusula `GRANTED BY`.

NOTA

Muchas aplicaciones de los proveedores contienen disposiciones para la asignación de privilegios del sistema tales como iniciar y detener la base de datos de los identificadores de autorización utilizando la instrucción `GRANT`. Ya que la sintaxis de estas variantes es totalmente específica a implementación, no se cubrirá aquí.

Se deben tener los privilegios necesarios en un objeto para otorgar privilegios en ese objeto. Si creó el objeto, entonces es el propietario, lo que significa que puede acceder por completo al objeto. (Todos los privilegios se le otorgan, incluyendo la habilidad para asignar privilegios a otros identificadores de autorización.)

Ahora echemos un vistazo a la cláusula `GRANT`. La cláusula incluye dos opciones: `ALL PRIVILEGES` y el marcador de posición `<lista de privilegios>`. Si utiliza las palabras clave `ALL PRIVILEGES`, se otorgan todos los privilegios disponibles a ese objeto de acuerdo con los privilegios que se otorgan sobre el objeto. Por ejemplo, asuma por un momento que se crea una tabla y es el propietario. Como resultado, automáticamente se le conceden los privilegios `SELECT`, `INSERT`, `UPDATE`, `DELETE`, `TRIGGER` y `REFERENCES`. (Éstos son los únicos privilegios que aplican a una tabla. Remítase a la tabla 6-1 para ver la lista de privilegios y de los objetos a los que se aplican.) También se le concede de forma automática la habilidad de asignar estos privilegios. En esta situación, si se utilizan las palabras clave `ALL PRIVILEGES`, se le otorgan estos seis privilegios a los identificadores de autorización en la instrucción `GRANT`.

Si decide no utilizar la opción `ALL PRIVILEGES`, debe entonces listar cada privilegio que debería aplicarse a los identificadores de usuario. Sin embargo, puede listar sólo los privilegios que se pueden aplicar al objeto específico. Por ejemplo, no se puede listar el privilegio `DELETE` si se otorga un privilegio en un dominio. También note que si se lista más de un privilegio, debe separar los nombres de los privilegios con comas.

La siguiente cláusula que veremos es la cláusula `ON`, que incluye dos marcadores de posición: `<tipo de objeto>` y `<nombre del objeto>`. El marcador de posición `<tipo de objeto>` simplemente se

refiere al tipo de objeto en el que está la concesión de permisos. SQL respalda los siguientes valores para el marcador de posición <tipo de objeto>:

- TABLE (incluye vistas)
- DOMAIN
- COLLATION
- CHARACTER SET
- TRANSLATION
- TYPE
- SEQUENCE
- Designación especial para rutinas invocadas SQL

Se requiere un valor para la designación <tipo de objeto>, a menos que el valor sea TABLE, en cuyo caso se puede dejar fuera. Si se proporciona el nombre de un objeto sin especificar un tipo, SQL asume que el valor del <tipo de objeto> es TABLE. Como se señala en la lista, la palabra clave TABLE también incluye vistas. Y por supuesto, no todas las aplicaciones respaldan todos los tipos de objetos incluidos en el estándar SQL, y algunos incluyen tipos de objetos que el estándar no cubre (lo cual conduce a la aplicación específica de las variaciones de SQL).

NOTA

En Oracle, la palabra clave TABLE se debe omitir. Por otra parte, en DB2 es opcional, pero ampliamente recomendado por el proveedor. SQL Server no respalda ninguna de las palabras clave de <tipo de objeto> del estándar SQL en la instrucción GRANT, pero en lugar de eso permite una clase de palabra clave seguida por un separador <::>; por lo tanto, en lugar de TABLE, se puede escribir OBJECT:: (OBJECT es la palabra clave para las bases de datos objetos tales como tablas y vistas). Es evidente que la lección aquí es que siempre consulte la documentación para la aplicación específica de SQL.

El marcador de posición <nombre del objeto> en la cláusula ON se refiere al nombre de un objeto específico. Siempre se requiere este valor.

La siguiente cláusula es la cláusula TO. Así como la cláusula GRANT, la cláusula TO tiene dos opciones: PUBLIC y el marcador de posición <lista de identificador de autorización>. Si se utiliza PUBLIC, todos los usuarios de la base de datos pueden acceder al objeto. Si se utiliza la opción <lista de identificador de autorización>, entonces se debe proporcionar el nombre de uno o más identificadores de autorización. Si se proporciona más de uno, deben estar separados por comas.

La última cláusula que se va a discutir es la cláusula WITH GRANT OPTION. Esta cláusula otorga a los identificadores de autorización el permiso para conceder cualquiera de los privilegios que se les otorgan en la instrucción GRANT. Por ejemplo, suponga que al identificador de usuario EmmaW se le otorga el privilegio SELECT en una de las tablas. Si se utiliza WITH GRANT OPTION, EmmaW podrá otorgar el privilegio SELECT a otro usuario. Si no se utiliza WITH GRANT OPTION, EmmaW no podrá otorgar el privilegio a otro usuario. Por cierto, la mayoría

de los expertos en seguridad recomiendan que nunca se utilice esta opción, ya que rápidamente se pierde el control sobre quién tiene qué privilegios.

Ahora que echamos un vistazo a la sintaxis, veamos algunos ejemplos. En el primer ejemplo, veamos la instrucción GRANT que otorga el privilegio SELECT al identificador de autorización PUBLIC. El privilegio se otorga a la vista llamada CD_DISPONIBLES, que enumera los CD que actualmente se tienen en existencia. Para otorgar el privilegio, utilice la siguiente instrucción:

```
GRANT SELECT ON TABLE CD_DISPONIBLES TO PUBLIC;
```

El privilegio SELECT permite a todos los usuarios de la base de datos (PUBLIC) ver los datos en la vista CD_DISPONIBLES. Sin embargo, ya que a PUBLIC no se le otorga ningún otro privilegio, los usuarios pueden ver los datos, pero no tomar ninguna acción. Además, ya que la cláusula WITH GRANT OPTION no se incluye en la instrucción, los usuarios no pueden asignar el privilegio SELECT a ningún otro usuario (que es un punto discutible en este caso porque todos pueden ya acceder a la vista CD_DISPONIBLES).

Ahora veamos otro ejemplo. Esta vez se otorgan los privilegios SELECT, UPDATE e INSERT a los roles VENTAS y CONTABILIDAD para que puedan acceder a la tabla INVENTARIO_CD:

```
GRANT SELECT, UPDATE, INSERT
ON TABLE INVENTARIO_CD
TO VENTAS, CONTABILIDAD WITH GRANT OPTION;
```

Observe que los privilegios se separan por comas, así como los roles. Como resultado de esta instrucción, los usuarios asociados con los roles VENTAS y CONTABILIDAD pueden ver, actualizar e insertar información en la tabla CD_INVENTARIO. Además, estos usuarios pueden asignar los privilegios SELECT, UPDATE e INSERT a otros usuarios que necesiten acceder a la tabla CD_INVENTARIO.

El siguiente ejemplo que se examina tiene una ligera variación de este último. Todo es igual, excepto que esta vez se especifica qué columna se puede actualizar:

```
GRANT SELECT, UPDATE(TITULO_CD), INSERT
ON TABLE INVENTARIO_CD
TO VENTAS, CONTABILIDAD WITH GRANT OPTION;
```

Observe que se puede añadir un nombre de columna después de un privilegio específico. Se pueden agregar nombres de columnas sólo a los privilegios SELECT, UPDATE e INSERT. Si se añade más de un nombre de columna, se deben separar por comas.

La instrucción GRANT en este ejemplo aún permite a los usuarios de VENTAS y CONTABILIDAD ver e insertar información en la tabla INVENTARIO_CD, pero sólo pueden actualizar valores en la columna TITULO_CD. No pueden actualizar ningún otro valor de columna en la tabla. Además, aunque puede aún asignar privilegios a otros usuarios, sólo pueden asignar el privilegio UPDATE en la columna TITULO_CD.

Echemos un vistazo a un ejemplo más que otorgue el privilegio SELECT al identificador de autorización PUBLIC:

```
GRANT SELECT (TITULO_CD, EN_EXISTENCIA) ON INVENTARIO_CD TO PUBLIC;
```

El identificador de autorización PUBLIC permite a todos los usuarios ver los datos en las columnas TITULO_CD y EN_EXISTENCIA de la tabla INVENTARIO_CD, pero no pueden ver ninguna otra información de esa tabla y no pueden modificar los datos en modo alguno. Observe en esta instrucción que la palabra clave TABLE no está incluida. Como se dijo antes, TABLE no se requiere.

La instrucción GRANT, cuando se utiliza conjuntamente con los privilegios disponibles y los identificadores de autorización, proporciona una base sólida para la seguridad de la base de datos. Sin embargo, cada aplicación de SQL es diferente con respecto a cómo se implementa y mantiene la seguridad. Por lo tanto, cuando se trata de cuestiones de seguridad, es importante que trabaje en estrecha colaboración con los administradores de red y base de datos y lea cuidadosamente la documentación del producto.

Revoque privilegios

Ahora que sabe cómo conceder privilegios a los identificadores de autorización, es hora de aprender cómo revocar esos privilegios. La instrucción que se utiliza para revocar privilegios es la instrucción REVOKE, como se muestra en la siguiente sintaxis:

```
REVOKE [ GRANT OPTION FOR ] { ALL PRIVILEGES | <lista de privilegios> }
ON <tipo de objeto> <nombre del objeto>
FROM { PUBLIC | <lista de identificador de autorización>
[GRANTED BY { CURRENT_USER | CURRENT_ROLE } ]
{RESTRICT | CASCADE}
```

Probablemente reconozca muchos de los elementos de la sintaxis de la instrucción GRANT o de otras instrucciones. De hecho, el único componente nuevo, que no sea la palabra clave REVOKE, es la cláusula GRANT OPTION FOR. Echemos un vistazo a la primera, ya que está al comienzo de la instrucción REVOKE. Esta cláusula aplica sólo cuando la cláusula WITH GRANT OPTION se utiliza en la instrucción GRANT. Si un privilegio se otorga con esta cláusula, se puede utilizar la cláusula GRANT OPTION FOR para eliminar ese permiso en particular. En caso de hacer uso de ésta, los privilegios se conservan, pero el usuario ya no puede otorgar esos privilegios a otros usuarios. Sin embargo, muy pocos productos RDBMS respaldan esta cláusula.

Olvide por un momento la cláusula GRANT OPTION FOR y veamos la cláusula REVOKE, que se utiliza para revocar todos los privilegios de un objeto (ALL PRIVILEGES) o sólo los privilegios definidos (<lista de privilegios>). Ambas opciones tienen el mismo significado que tuvieron en la instrucción GRANT; puede utilizar ALL PRIVILEGES o puede listar cada privilegio separado por una coma.

Las cláusulas ON y GRANTED BY en la instrucción REVOKE son exactamente las mismas que las cláusulas ON y GRANTED BY en la instrucción GRANT. Para la cláusula ON, se deben especificar valores para los marcadores de posición <tipo de objeto> y <nombre del objeto>; sin embargo, si el valor del <tipo de objeto> es TABLE, entonces se puede dejar fuera (y como antes, se debe omitir en Oracle y SQL Server). Para la cláusula GRANTED BY, asumiendo que

el RDBMS lo respalda (la mayoría no lo hace), puede escoger una de dos opciones (CURRENT_USER o CURRENT_ROLE).

La cláusula FROM en la instrucción REVOKE se puede comparar también a la instrucción GRANT. La única diferencia es que en la instrucción GRANT se utiliza la palabra clave TO, pero en la instrucción REVOKE se usa la palabra clave FROM. En ambos casos, se debe utilizar PUBLIC como identificador de autorización, o se deben enumerar los identificadores de usuario y los nombres de rol.

Los últimos elementos de la instrucción a analizar son las palabras clave RESTRICT y CASCADE. Probablemente recuerde estas palabras clave de los capítulos 2, 3 y 4. Si se especifica RESTRICT, el privilegio no se revoca si se pasa a otros usuarios (en otras palabras, si hay algún privilegio dependiente). (Esto significaría que WITH GRANT OPTION se utilizó en la instrucción GRANT y que el identificador de autorización que otorgó el privilegio entonces concedió el privilegio a alguien más.) Si se especifica CASCADE, el privilegio se revoca así como cualquiera de los privilegios que se pasan a otros usuarios.

NOTA

Las aplicaciones de los proveedores varían. En Oracle, CASCADE debe especificarse como CASCADE CONSTRAINTS. En ambos, Oracle y SQL Server, no se puede especificar RESTRICT, sino que más bien es el comportamiento predeterminado cuando no se especifica CASCADE. MySQL no permite que se especifique ninguno.

Ahora veamos algunos ejemplos de revocación de privilegios. La siguiente instrucción revoca el privilegio SELECT que se otorga al identificador de autorización PUBLIC en la vista CD_DISPONIBLES:

```
REVOKE SELECT ON TABLE CD_DISPONIBLES FROM PUBLIC CASCADE;
```

Como se puede observar, esta instrucción es muy similar a la instrucción GRANT. Se deben identificar los privilegios, los identificadores de autorización y el objeto. Además, se debe especificar RESTRICT o CASCADE.

El siguiente ejemplo se basa en privilegios que se otorgan a la tabla llamada INVENTARIO_CD. A los roles VENTAS y CONTABILIDAD se les otorgan los siguientes privilegios en esta tabla: GRANT, SELECT e INSERT. Para revocar esos privilegios, utilice la siguiente instrucción REVOKE:

```
REVOKE SELECT, UPDATE, INSERT ON TABLE INVENTARIO_CD  
FROM VENTAS, CONTABILIDAD CASCADE;
```

Observe que sólo tiene que especificar los privilegios que quiere revocar, el nombre de los objetos y el nombre de los identificadores de autorización. Sin embargo, ya que se están revocando todos los privilegios que se otorgaron, se puede simplificar la instrucción utilizando las palabras clave ALL PRIVILEGES, como se muestra en el siguiente ejemplo:

```
REVOKE ALL PRIVILEGES ON TABLE INVENTARIO_CD  
FROM VENTAS, CONTABILIDAD CASCADE;
```

Si no desea revocar todos los privilegios, pero en vez de eso quiere revocar sólo los privilegios UPDATE e INSERT, se pueden especificar sólo estos privilegios, como se muestra en el siguiente ejemplo:

```
REVOKE UPDATE, INSERT ON TABLE INVENTARIO_CD
FROM VENTAS, CONTABILIDAD CASCADE;
```

También se puede elegir revocar privilegios para sólo uno de los nombres de rol, en lugar de ambos. Además, se puede utilizar la palabra clave RESTRICT en lugar de CASCADE.

Ahora suponga que se otorgan los mismos privilegios como en el ejemplo anterior, pero además de éstos, se especifica WITH GRANT OPTION cuando se conceden los privilegios. Si se quiere revocar sólo la habilidad de los roles Ventas y Contabilidad de otorgar privilegios a otros usuarios, se puede utilizar la siguiente instrucción:

```
REVOKE GRANT OPTION FOR ALL PRIVILEGES ON INVENTARIO_CD
FROM VENTAS, CONTABILIDAD CASCADE;
```

Esta instrucción revoca sólo la habilidad de conceder privilegios; los roles Ventas y Contabilidad aún pueden acceder a la tabla INVENTARIO_CD. Si se desea revocar todos los privilegios, se tiene que ejecutar esta instrucción sin la cláusula GRANT OPTION FOR. Observe en esta instrucción que la palabra clave TABLE no se usa antes del nombre de la tabla. La instrucción REVOKE, como la instrucción GRANT, no requiere la palabra clave TABLE cuando se especifica una tabla o vista.

NOTA

Las diferencias de implementación del proveedor indicadas con el uso de la palabra clave TABLE en la instrucción GRANT (anteriormente en este capítulo) aplican de la misma forma a la instrucción REVOKE. En general, los proveedores respaldan la sintaxis idéntica entre las instrucciones GRANT y REVOKE, exceptuando que TO en la instrucción GRANT se convierte en FROM en la instrucción REVOKE.

Otorgue y revoque roles

Ahora que sabe cómo crear y eliminar roles y conceder y revocar privilegios, veamos la concesión y revocación de roles. Empezaremos con la concesión de roles. Para otorgar un rol, se debe utilizar la instrucción GRANT para asignar uno o más nombres de rol o uno o más identificadores de autorización, como se muestra en la siguiente sintaxis:

```
GRANT <lista de nombres de rol>
TO { PUBLIC | <lista de identificador de autorización> } [ WITH ADMIN OPTION ]
[ GRANTED BY { CURRENT_USER | CURRENT_ROLE } ]
```

Por ahora, la mayor parte de esta sintaxis debe lucir bastante familiar, excepto por algunas variaciones. La cláusula GRANT permite especificar un listado de uno o más nombres de rol. Si se especifica más de un nombre, se deben separar por comas. La cláusula TO permite especificar uno o más identificadores de autorización. Una vez más, si hay más de uno, se deben separar por comas. También se puede especificar el identificador de autorización PUBLIC para otorgar un rol a todos los usuarios de la base de datos. La cláusula WITH ADMIN OPTION, que es opcional,

permite a los identificadores de autorización otorgar un rol a otros usuarios. Y la cláusula GRANTED BY, que también es opcional (y sólo se soporta en algunos productos RDBMS), se utiliza en los casos raros cuando el par identificador de usuario/nombre de rol no contiene un valor nulo.

Echemos un vistazo a otro ejemplo. Suponga que tiene que crear un rol llamado ADMINISTRADORES y desea asignar ese rol a un identificador de usuario llamado LindaN. Se utiliza la siguiente sintaxis:

```
GRANT ADMINISTRADORES TO LindaN;
```

Ahora suponga que quiere darle a LindaN la habilidad de otorgar el rol ADMINISTRADORES a otros usuarios. Para hacer esto, simplemente añada la cláusula WITH ADMIN OPTION, como en el siguiente ejemplo:

```
GRANT ADMINISTRADORES TO LindaN WITH ADMIN OPTION;
```

También se pueden conceder múltiples roles a múltiples identificadores de usuario. Los identificadores de usuario pueden ser identificadores de usuarios u otros nombres de rol. En el siguiente ejemplo se otorgan los roles ADMINISTRADORES y CONTABILIDAD al identificador de usuario LindaN y el nombre de rol MARKETING:

```
GRANT ADMINISTRADORES, CONTABILIDAD TO LindaN, MARKETING WITH ADMIN OPTION;
```

Ahora que sabe cómo otorgar roles a los identificadores de autorización, es hora de aprender cómo revocar esos roles.

Revoque roles

La revocación de roles es muy similar a la revocación de privilegios. La instrucción que se utiliza para revocar privilegios es la instrucción REVOKE, como se muestra en la siguiente sintaxis:

```
REVOKE [ ADMIN OPTION FOR ] <lista de nombres de rol>  
FROM { PUBLIC | <lista de identificador de autorización> }  
[ GRANTED BY { CURRENT_USER | CURRENT_ROLE } ]  
{ RESTRICT | CASCADE }
```

Como se puede observar, no hay nada nuevo en la sintaxis excepto por la cláusula ADMIN OPTION FOR, que es similar a la cláusula GRANT OPTION FOR utilizada cuando se revocan privilegios. La cláusula permite revocar la habilidad de asignar roles a otros usuarios, sin revocar el propio rol.

NOTA

Las mismas diferencias mencionadas para Oracle, SQL Server y MySQL en cuanto a la revocación de privilegios aplican al uso de la instrucción REVOKE.

Veamos un ejemplo de la revocación de un rol. Suponga que otorga el rol ADMINISTRADORES al identificador de autorización LindaN. Se puede revocar ese rol utilizando la siguiente instrucción REVOKE:

```
REVOKE ADMINISTRADORES FROM LindaN CASCADE;
```

Si se otorgan los roles ADMINISTRADORES y CONTABILIDAD a LindaN y el rol MARKETING, la instrucción REVOKE sería como la siguiente:

```
REVOKE ADMINISTRADORES, CONTABILIDAD FROM LindaN, MARKETING CASCADE;
```

Ahora que hemos visto cómo otorgar y revocar roles, se puede ver cómo esto es similar a la concesión y revocación de privilegios. Una vez más, debo destacar que no todas las aplicaciones son similares con respecto a cómo otorgan y revocan privilegios y roles, así que asegúrese de revisar la documentación del producto y trabaje en estrecha colaboración con el administrador de bases de datos.

Pruebe esto 6-1 Gestión de roles y privilegios

En este ejercicio se crearán dos roles en la base de datos INVENTARIO, se otorgarán privilegios al identificador de autorización PUBLIC y a uno de los roles que se crea, se concederá uno de los roles al otro rol, y luego se revocarán todos los privilegios y roles. Finalmente, se eliminarán los dos roles que se crearon. La habilidad de seguir todos los pasos en este ejercicio dependerá del tipo de instrucciones relacionadas con la seguridad respaldada en la aplicación de SQL que esté utilizando. Sin embargo, el ejercicio está diseñado para que cualquiera de los roles que se creen o cualquiera de los privilegios que se asignen sean eliminados al final del ejercicio. No se utilizarán estos roles para ninguno de los siguientes ejercicios. Si por cualquier motivo este ejercicio afectara la seguridad del sistema en el que está trabajando, debería discutir este ejercicio con el administrador de la base de datos u omitirlo por completo. Puede descargar el archivo Try_This_06.txt, que contiene las instrucciones SQL utilizadas en este ejercicio (en inglés).

Paso a paso

1. Abra la aplicación de cliente de su RDBMS y conéctese a la base de datos INVENTARIO.
2. Lo primero que se hace es crear el rol MRKT. Introduzca y ejecute la siguiente instrucción SQL:

```
CREATE ROLE MRKT;
```
3. Después se crea el rol PERSONAL_VENTAS. Introduzca y ejecute la siguiente instrucción SQL:

```
CREATE ROLE PERSONAL_VENTAS;
```
4. Ahora se otorga el privilegio SELECT en la vista CD_EN_EXISTENCIA. El privilegio se asigna al identificador de autorización PUBLIC. Introduzca y ejecute la siguiente instrucción SQL:

```
GRANT SELECT ON CD_EN_EXISTENCIA TO PUBLIC;
```
5. El siguiente privilegio que se otorga es el rol PERSONAL_VENTAS que se creó en el paso 3. Se otorgan los privilegios SELECT, INSERT y UPDATE en la tabla DISCOS_COMPACTOS. Para el privilegio UPDATE se especifica la columna TITULO_CD. También se le permite al rol

(continúa)

STAFF_VENTAS otorgar estos privilegios a otros usuarios. Introduzca y ejecute la siguiente instrucción SQL:

```
GRANT SELECT, INSERT, UPDATE(TITULO_CD) ON DISCOS_COMPACTOS  
TO PERSONAL_VENTAS WITH GRANT OPTION;
```

- 6.** Ahora se otorga el rol PERSONAL_VENTAS al rol MRKT. Introduzca y ejecute la siguiente instrucción SQL:

```
GRANT PERSONAL_VENTAS TO MRKT;
```

- 7.** El siguiente paso es revocar el privilegio SELECT que se otorgó al identificador de autorización PUBLIC. Introduzca y ejecute la siguiente instrucción SQL:

```
REVOKE SELECT ON CD_EN_EXISTENCIA FROM PUBLIC CASCADE;
```

- 8.** Ahora se revocan los privilegios que se concedieron al rol PERSONAL_VENTAS. Ya que se revocan todos los privilegios, se puede utilizar la palabra clave ALL PRIVILEGES. También debe asegurarse de revocar cualquier privilegio dependiente; por lo tanto, se utiliza la palabra clave CASCADE. Introduzca y ejecute la siguiente instrucción SQL:

```
REVOKE ALL PRIVILEGES ON DISCOS_COMPACTOS FROM PERSONAL_VENTAS CASCADE;
```

- 9.** Ahora se puede revocar el rol PERSONAL_VENTAS del rol MRKT. Introduzca y ejecute la siguiente instrucción SQL:

```
REVOKE PERSONAL_VENTAS FROM MRKT CASCADE;
```

- 10.** El siguiente paso es eliminar el rol MRKT. Introduzca y ejecute la siguiente instrucción SQL:

```
DROP ROLE MRKT;
```

- 11.** Finalmente, se requiere eliminar el rol PERSONAL_VENTAS. Introduzca y ejecute la siguiente instrucción SQL:

```
DROP ROLE PERSONAL_VENTAS;
```

- 12.** Cierre la aplicación clientes.

Resumen de Pruebe esto

La base de datos INVENTARIO debe estar compuesta de la misma manera como antes de comenzar este ejercicio. Los permisos y roles concedidos se revocaron, y los roles que se crearon se eliminaron. De esta manera, no tendrá que preocuparse por las consideraciones de seguridad para otros ejercicios. Para el resto de los ejercicios en el libro, se continuará trabajando en el mismo contexto de seguridad en el que se ha trabajado para este ejercicio y para todos los ejercicios anteriores al presente.

Autoexamen Capítulo 6

1. ¿Cuál es la diferencia entre un identificador de usuario y un nombre de rol?
2. ¿Cuál es el nombre del identificador de autorización especial que otorga el acceso a todos los usuarios de la base de datos?
3. Cada _____ se asocia con un identificador de usuario y un nombre de rol.
4. ¿Con cuál de los siguientes se asocia una sesión SQL?
 - A Privilegio
 - B Identificador de autorización
 - C PUBLIC
 - D Nombre de rol
5. Cuando se establece por primera vez una sesión SQL, el identificador de usuario siempre es el _____.
6. ¿Cuál es el valor del nombre de rol actual cuando se establece por primera vez una sesión SQL?
7. ¿Qué es un identificador de autorización?
8. ¿Cuáles son los dos tipos de identificadores de autorización que respalda SQL?
9. ¿Qué privilegios se le deben otorgar a un objeto si desea permitir que un identificador de autorización consulte los datos de ese objeto?
10. Se establece una sesión SQL con la base de datos. El identificador de usuario actual es EthanW. El nombre de rol actual es nulo. ¿Cuál es el identificador de autorización actual?
11. ¿En qué objetos de esquema se pueden definir privilegios de acceso?
12. ¿En qué tipos de objetos de base de datos se puede asignar el privilegio DELETE?
 - A Tablas
 - B Vistas
 - C Columnas
 - D Dominios
13. ¿En qué tipos de objetos de base de datos se puede asignar el privilegio TRIGGER?
 - A Tablas
 - B Vistas
 - C Columnas
 - D Dominios
14. Se crea un rol llamado CONTABILIDAD. ¿Qué instrucción SQL deberá utilizarse?

- 15.** Se otorgan todos los privilegios de la vista NOMBRES_CD a todos los que utilizan la base de datos. ¿Qué instrucción SQL deberá utilizarse?
- 16.** Se otorga el privilegio SELECT al rol EMPLEADO_VENTAS a una tabla de la base de datos. Se desea que el rol EMPLEADO_VENTAS sea capaz de asignar el privilegio SELECT a otros usuarios. ¿Qué cláusula deberá incluirse en la instrucción GRANT?
- 17.** Se desea conceder el rol ACCT al usuario autorizado MaxN. No se quiere que el usuario pueda otorgar el rol a otros usuarios. ¿Qué instrucción SQL deberá utilizarse para otorgar el rol?

Parte II

Acceso y modificación de datos



Capítulo 7

Consulta de datos de SQL

Habilidades y conceptos clave

- Utilice la instrucción SELECT para la recuperación de datos
 - Utilice la cláusula WHERE para definir condiciones de búsqueda
 - Utilice la cláusula GROUP BY para agrupar los resultados de una consulta
 - Utilice la cláusula HAVING para especificar un grupo de condiciones de búsqueda
 - Utilice la cláusula ORDER BY para ordenar los resultados de una consulta
-

Una vez que se crearon los objetos en una base de datos y las tablas base están pobladas de datos, se pueden presentar las consultas que permiten recuperar información específica de la base de datos. Estas consultas, que suelen adoptar la forma de instrucciones SELECT, pueden variar en complejidad desde una simple instrucción que devuelve todas las columnas de una tabla a una instrucción que una múltiples tablas, calcule valores y defina condiciones de búsqueda que restrinja exactamente qué filas de datos se deben devolver. La instrucción SELECT se compone de una serie de cláusulas flexibles que juntas determinan qué datos se recuperarán. En este capítulo aprenderá cómo utilizar cada una de estas cláusulas para realizar tareas básicas de recuperación de datos, definir condiciones de búsqueda, agrupar los resultados de una consulta, especificar un grupo de condiciones de búsqueda y ordenar los resultados de una búsqueda.

Utilice la instrucción SELECT para la recuperación de datos

En el capítulo 5, cuando se analizó acerca de las vistas, se presentó la instrucción SELECT. Como podrá recordar, la instrucción SELECT permite formar consultas intrincadas que pueden devolver exactamente el tipo de datos que se desea recuperar. Es una de las instrucciones más comunes que utilizará como programador de SQL, y también es una de las instrucciones más flexible y amplia en el estándar SQL.

La instrucción SELECT es una expresión de consulta que comienza con la palabra clave SELECT e incluye una serie de elementos que forman la expresión. La sintaxis básica para la instrucción SELECT puede dividirse en varias cláusulas específicas, cada una de las cuales ayuda a refinar la consulta para que sólo se devuelvan los datos requeridos. La sintaxis para la instrucción SELECT puede mostrarse como sigue:

```
SELECT [ DISTINCT | ALL ] { * | <selección de lista> }  
FROM <tabla de referencia> [ { , <tabla de referencia> }... ]  
[ WHERE <condición de búsqueda> ]  
[ GROUP BY <especificación de agrupación> ]  
[ HAVING <condición de búsqueda> ]  
[ ORDER BY <condición de orden> ]
```

Como se puede observar, las únicas cláusulas requeridas son la cláusula `SELECT` y la cláusula `FROM`. Todas las demás cláusulas son opcionales.

Las cláusulas `FROM`, `WHERE`, `GROUP BY` y `HAVING` hacen referencia como expresiones de tabla. Esta porción de la instrucción `SELECT` siempre se evalúa en el orden indicado en la sintaxis. El resultado de esa evaluación es una tabla virtual que se utiliza en la evaluación posterior. En otras palabras, los resultados de la primera cláusula evaluada se utilizan en la siguiente cláusula. Los resultados de esa cláusula se utilizan en la siguiente cláusula, hasta que cada cláusula en la expresión de la tabla se evalúa. Por ejemplo, la primera cláusula a evaluar en la instrucción `SELECT` es la cláusula `FROM`. Ya que esta cláusula es necesaria, siempre es la primera cláusula evaluada. Los resultados de la cláusula `FROM` se utilizan en la cláusula `WHERE`, si se especifica la cláusula `WHERE`. Si no se especifica la cláusula, entonces los resultados de la cláusula `FROM` se utilizan en la siguiente cláusula especificada, ya sea la cláusula `GROUP BY` o la cláusula `HAVING`. Una vez que se evalúa la cláusula final en la expresión de tabla, los resultados se utilizan en la cláusula `SELECT`. Después de evaluar la cláusula `SELECT`, se evalúa la cláusula `ORDER BY`.

Para resumir todo esto, las cláusulas de la instrucción `SELECT` se aplican en el siguiente orden:

- Cláusula `FROM`
- Cláusula `WHERE` (opcional)
- Cláusula `GROUP BY` (opcional)
- Cláusula `HAVING` (opcional)
- Cláusula `SELECT`
- Cláusula `ORDER BY` (opcional)

Es importante tener un entendimiento básico del orden de evaluación al crear instrucciones `SELECT` más complejas, especialmente cuando se trabaja con uniones y subconsultas (discutidas en los capítulos 11 y 12, respectivamente). Este entendimiento es también útil cuando se analiza cada cláusula individualmente, ya que explica cómo se relaciona una cláusula con otras cláusulas. Como resultado, es una buena idea mantener este orden de evaluación en mente a lo largo de este capítulo y en los capítulos siguientes que se basan en diversos aspectos de la instrucción `SELECT`.

La cláusula `SELECT` y la cláusula `FROM`

Ahora que tiene una visión general básica de cómo la instrucción `SELECT` se ejecuta, echemos un vistazo más de cerca a la cláusula `SELECT` y la cláusula `FROM`, las dos cláusulas requeridas en la instrucción. Se analizarán las otras cláusulas en secciones separadas durante el resto del capítulo.

Empecemos con la cláusula `SELECT`. La cláusula `SELECT` incluye las palabras clave `DISTINCT` y `ALL`. La palabra clave `DISTINCT` se utiliza si se desean eliminar filas duplicadas de los resultados de la consulta, y la palabra clave `ALL` se utiliza si se quieren devolver todas las filas de los resultados de una consulta. Por ejemplo, suponga que la base de datos incluye una tabla llamada `CD_INTERPRETE`. La tabla incluye las columnas `NOMBRE_INTERPRETE` y `NOMBRE_CD`. Ya que un `CD` puede incluir más de un intérprete, el nombre del `CD` puede aparecer más de una vez en la tabla. Ahora suponga que desea consultar la tabla sólo por el nombre de los `CD`, pero

no quiere los nombres repetidos. Puede utilizar la palabra clave `DISTINCT` para asegurar que la consulta devuelva el nombre de cada CD sólo una vez, o puede utilizar la palabra clave `ALL` para especificar que se devuelvan todas las filas, aun si hay filas duplicadas. Si no se especifica ninguna de las palabras clave, se toma la palabra clave `ALL`.

Pregunta al experto

P: Se afirma que se puede utilizar un asterisco para incluir todas las columnas en el resultado de una consulta. ¿Presenta esto algún problema si el número de columnas cambia?

R: Sí, esto puede presentar un problema. De hecho, generalmente se recomienda que se utilice el asterisco sólo cuando se accede a una base de datos SQL a través de una invocación directa. Si se utiliza el asterisco en SQL incrustado y el número de columnas cambia, puede encontrar que la aplicación ya no responda correctamente, ya que el programa de aplicación se codificó para esperar una respuesta específica. Si se utiliza un asterisco y se espera que columnas específicas se devuelvan, entonces puede encontrarse con una serie de sorpresas si la base de datos ha cambiado. Por este motivo, se debe evitar el asterisco a menos que se invoque directamente la instrucción `SELECT`. Sin embargo, en el caso de la invocación directa, el asterisco es una manera práctica de devolver todas las columnas sin tener que especificar el nombre de cada una. De hecho, muchos de los ejemplos en el capítulo utilizan el asterisco para evitar el tener que repetir los nombres de las columnas innecesariamente.

Además de las palabras clave `DISTINCT` y `ALL`, la cláusula `SELECT` incluye el asterisco (*) y el marcador de posición <selección de lista>. Se debe especificar una de estas opciones en la cláusula. Si se especifica el asterisco, todas las columnas aplicables se incluyen en el resultado de la consulta.

Si no se especifica el asterisco en la cláusula `SELECT`, se debe especificar cada columna tal como se deriva del origen. El marcador de posición <selección de lista> se puede desglosar en la siguiente sintaxis:

```
<columna derivada> [ [AS] <nombre de la columna> ]
[ { , <columna derivada> [ [AS] <nombre de la columna> } ... ]
```

Echemos un vistazo a la primera línea de la sintaxis. (La segunda línea no es más que una repetición, tantas veces como sea necesario, de la primera línea.) El marcador de posición <columna derivada> en la mayoría de los casos se refiere al nombre de la columna en la tabla origen. Si más de una columna se especifica, entonces se deben separar por comas. Sin embargo, el marcador de posición <columna derivada> también puede referirse a una columna o conjunto de columnas que son de alguna manera parte de una expresión. Por ejemplo, en el capítulo 5 se analizó la función `AVG`, que saca el promedio de los valores en una columna específica. El ejemplo que se muestra en este capítulo utiliza la instrucción `SELECT` para consultar datos de la tabla `COMISIONES_EMPLEADO`, que enumera la cantidad total de comisiones hechas por cada empleado durante un

periodo de tres años. La instrucción SELECT promedia los valores en tres diferentes columnas, como se muestra en la siguiente instrucción SELECT:

```
SELECT AVG(AÑO_1999) , AVG(AÑO_2000) , AVG(AÑO_2001)
FROM COMISIONES_EMPLEADO;
```

En este caso, hay tres expresiones que se utilizan para el marcador de posición <columna derivada>: AVG(AÑO_1999), AVG(AÑO_2000) y AVG(AÑO_2001). Observe que la expresión de cada columna derivada se separa por una coma, como sería el caso si cada valor fuera simplemente un nombre de columna. El siguiente ejemplo muestra la misma instrucción SELECT como en el ejemplo anterior, excepto que se utilizan sólo los nombres de las columnas como columnas derivadas:

```
SELECT AÑO_1999, AÑO_2000, AÑO_2001
FROM COMISIONES_EMPLEADO;
```

Si se fuera a ejecutar esta instrucción SELECT, la consulta devolvería todos los valores en las tres columnas, en lugar de promediar esos valores.

La cláusula SELECT también permite proporcionar un nombre de columna para cada columna derivada. Para hacer esto, se añade la palabra clave AS y el nuevo nombre de columna después de la columna derivada, como se muestra en el siguiente ejemplo:

```
SELECT AVG(AÑO_1999) AS PROMEDIO_1999
FROM COMISIONES_EMPLEADO;
```

En esta instrucción SELECT, el valor que se devuelve de la columna AÑO_1999 se coloca en la columna llamada PROMEDIO_1999. Éste es el nombre de la columna que se devuelve como parte de una tabla virtual en los resultados de la consulta. Si no se especifica la subcláusula AS, el nombre de columna en la tabla virtual es el mismo que el nombre de columna en la tabla origen. Si un nombre de columna no se puede heredar naturalmente (por ejemplo, cuando se agregan dos valores de columna), se debe utilizar la subcláusula AS.

Observe que en los ejemplos anteriores la cláusula FROM se utiliza para especificar la tabla (COMISIONES_EMPLEADO) que contiene las columnas mencionadas en la cláusula SELECT. La cláusula FROM incluye la palabra clave FROM y una o más tablas de referencia. Si hay varias tablas de referencia, deben separarse utilizando comas. En la mayoría de los casos, la tabla de referencia es el nombre de una tabla o de tablas unidas, aunque puede también ser un tipo de subconsultas. Se discutirá la unión de tablas en el capítulo 11 y las subconsultas en el capítulo 12. Para este capítulo, la cláusula FROM se utiliza principalmente para hacer referencia a nombres de tablas, como se define la cláusula en los dos ejemplos anteriores (donde <tabla de referencia> es igual a COMISIONES_EMPLEADO).

Juntas las cláusulas SELECT y FROM forman las bases para la instrucción SELECT, que puede ser tan simple como consultar cada fila y cada columna de una tabla, como se muestra en el siguiente ejemplo:

```
SELECT * FROM INTERPRETES;
```

En esta instrucción, se especifica que cada columna de la tabla INTERPRETES se debe devolver. Además, cada fila será devuelta, ya que no se especificaron otras cláusulas.

ID_INTERPRETE: INT	NOMBRE_INTERPRETE: VARCHAR(60)	LUGAR_DE_NACIMIENTO: VARCHAR(60)
2001	Jennifer Warnes	Seattle, Washington, USA
2002	Joni Mitchell	Fort MacLeod, Alberta, Canada
2003	William Ackerman	Alemania
2004	Kitaro	Toyohashi, Japon
2005	Bing Crosby	Tacoma, Washington, Estados Unidos
2006	Patsy Cline	Winchester, Virginia, Estados Unidos
2007	Jose Carreras	Barcelona, España
2008	Luciano Pavarotti	Modena, Italia
2009	Placido Domingo	Madrid, España

Figura 7-1 Las columnas ID_INTERPRETE, NOMBRE_INTERPRETE y LUGAR_DE_NACIMIENTO de la tabla INTERPRETES.

Echemos un vistazo más de cerca de esto. La tabla INTERPRETES incluye las columnas ID_INTERPRETE, NOMBRE_INTERPRETE y LUGAR_DE_NACIMIENTO, como se muestra en la figura 7-1.

Si se ejecuta la instrucción SELECT mostrada en el ejemplo anterior, los resultados de la consulta serían similares a los siguientes:

ID_INTERPRETE	NOMBRE_INTERPRETE	LUGAR_DE_NACIMIENTO
-----	-----	-----
2001	Jennifer Warnes	Seattle, Washington, Estados Unidos
2002	Joni Mitchell	Fort MacLeod, Alberta, Canadá
2003	William Ackerman	Alemania
2004	Kitaro	Toyohashi, Japón
2005	Bing Crosby	Tacoma, Washington, Estados Unidos
2006	Patsy Cline	Winchester, Virginia, Estados Unidos
2007	Jose Carreras	Barcelona, España
2008	Luciano Pavarotti	Modena, Italia
2009	Placido Domingo	Madrid, España

Observe que cada fila y cada columna se devuelven en los resultados de la consulta. Si utiliza el asterisco en la cláusula SELECT, no se tienen que especificar los nombres de columnas.

Ahora suponga que desea devolver sólo las columnas NOMBRE_INTERPRETE y LUGAR_DE_NACIMIENTO. Se puede modificar la instrucción SELECT para verse como la siguiente:

```
SELECT NOMBRE_INTERPRETE AS NAME, LUGAR_DE_NACIMIENTO
FROM INTERPRETES;
```

Los resultados de la consulta ahora contienen sólo dos columnas, como se muestran a continuación:

NOMBRE	LUGAR_DE_NACIMIENTO
Jennifer Warnes	Seattle, Washington, Estados Unidos
Joni Mitchell	Fort MacLeod, Alberta, Canadá
William Ackerman	Alemania
Kitaro	Toyohashi, Japón
Bing Crosby	Tacoma, Washington, Estados Unidos
Patsy Cline	Winchester, Virginia, Estados Unidos
Jose Carreras	Barcelona, España
Luciano Pavarotti	Modena, Italia
Placido Domingo	Madrid, España

Observe que el nombre de la primera columna es NOMBRE, en lugar de NOMBRE_INTERPRETE. Esto es debido a que la subcláusula AS (especificada en NOMBRE) se define como parte de la columna derivada NOMBRE_INTERPRETE. Si se fuera a especificar la palabra clave DISTINCT en esta situación particular, aún recibirían el mismo número de filas, aunque podrían no ser devueltas en el mismo orden en que estaban cuando no se utilizaba la palabra clave, dependiendo de la aplicación SQL. La razón por la que la palabra clave DISTINCT no haría ninguna diferencia en los resultados de la consulta es porque no hay filas duplicadas en la tabla. Sin embargo, utilizar la palabra clave DISTINCT puede afectar la ejecución, particularmente si el RDBMS tiene que ordenar a través de un gran número de filas, así que asegúrese de utilizar la palabra clave sólo cuando sea necesario.

Ahora veamos un ejemplo que utilice la palabra clave DISTINCT. Suponga que la base de datos incluye una tabla que empareje a los intérpretes con los tipos de música, como se muestra en la figura 7-2.

NOMBRE_INTERPRETE: VARCHAR(60)	TIPO_INTERPRETE: VARCHAR(10)
Jennifer Warnes	Folk
Jennifer Warnes	Pop
Joni Mitchell	Pop
Joni Mitchell	Folk
Joni Mitchell	Jazz
William Ackerman	New Age
Kitaro	New Age
Kitaro	International

Figura 7-2 Columnas NOMBRE_INTERPRETE y TIPO_INTERPRETE de la tabla TIPO_INTERPRETE.

Si la instrucción `SELECT` incluye ambas (todas) columnas en la cláusula `SELECT`, como se muestra en el siguiente ejemplo, la consulta devolvería todas las filas:

```
SELECT * FROM TIPO_INTERPRETE;
```

No importa si se especifica la palabra clave `DISTINCT` en este caso, ya que los resultados de la consulta no incluyen filas duplicadas. Los resultados serían los mismos si se incluyera la palabra clave `ALL`, en lugar de `DISTINCT`, o si no se especifican ninguno de los dos calificadores. En ambos casos, los resultados de la consulta incluyen la misma información que se muestra en la tabla de la figura 7-2.

Ahora echemos un vistazo a la misma instrucción, sólo que esta vez se especifica la palabra clave `DISTINCT` y sólo una de las dos columnas:

```
SELECT DISTINCT NOMBRE_INTERPRETE  
FROM TIPO_INTERPRETE;
```

Observe que esta instrucción incluye sólo la columna `NOMBRE_INTERPRETE`, que incluye valores duplicados. Mediante el uso de la palabra clave `DISTINCT`, los resultados de la consulta incluirían sólo un ejemplo de cada uno de los valores. Si se ejecuta la instrucción `SELECT` en el ejemplo anterior, los resultados de la consulta serían similares a los siguientes:

```
NOMBRE_INTERPRETE  
-----  
Jennifer Warnes  
Joni Mitchell  
Kitaro  
William Ackerman
```

Aunque hay siete filas en la tabla `TIPO_INTERPRETE`, sólo se devuelven cuatro filas, ya que existen sólo cuatro valores únicos en la columna `NOMBRE_INTERPRETE` y los demás valores son duplicados.

Como se puede observar, las cláusulas `SELECT` y `FROM` son bastante sencillas, al menos en este nivel de codificación. Una vez que lleguemos a estructuras más complejas, encontrará que ambas cláusulas pueden ser más complicadas. Sin embargo, lo importante para recordar ahora es que estas cláusulas actúan como base para el resto de la instrucción `SELECT`. En términos de ejecución, la instrucción `SELECT`, para todos los efectos prácticos, comienza con la cláusula `FROM` y termina con la cláusula `SELECT`. (La cláusula `ORDER BY` se utiliza principalmente para mostrar los objetivos, y de hecho no afecta la información devuelta. La cláusula `ORDER BY` se discute con mayor detalle en la sección “Utilice la cláusula `ORDER BY` para ordenar los resultados de una consulta”, más adelante en este capítulo.)

Utilice la cláusula `WHERE` para definir condiciones de búsqueda

La siguiente cláusula en la instrucción `SELECT` es la cláusula `WHERE`. La cláusula `WHERE` toma los valores devueltos por la cláusula `FROM` (en la tabla virtual) y aplica la condición de búsqueda que se define en la cláusula `WHERE`. La cláusula `WHERE` actúa como filtro sobre los resultados devueltos por la cláusula `FROM`. Cada fila se evalúa contra las condiciones de búsqueda.

da. Las filas que se evalúan como verdaderas se devuelven como parte del resultado de la consulta. Aquellas que se evalúan como desconocidas o falsas no se incluyen en los resultados.

Para una mejor comprensión de cómo se evalúa cada fila, echemos un vistazo de cerca al marcador de posición <condición de búsqueda>. La condición de búsqueda se compone de uno o más predicados que se utilizan para poner a prueba el contenido devuelto por la cláusula FROM. Un *predicado* es una expresión de SQL que define un hecho acerca de cualquier fila devuelta por la instrucción SELECT. Ya se vieron ejemplos de predicados en los capítulos 4 y 5. Por ejemplo, un ejemplo de definición de una vista (en el capítulo 5) incluye la siguiente instrucción SELECT:

```
SELECT TITULO_CD, DERECHOSDEAUTOR, EN_EXISTENCIA
FROM INVENTARIO_DISCO_COMPACTO
WHERE DERECHOSDEAUTOR > 1989 Y DERECHOSDEAUTOR < 2000;
```

Esta instrucción consulta tres columnas en la tabla INVENTARIO_DISCO_COMPACTO. La cláusula SELECT especifica las columnas que se devuelven, y la cláusula FROM especifica la tabla origen. La cláusula WHERE determina qué filas (basado en la cláusula FROM) se incluyen en los resultados. En este caso, la cláusula WHERE contiene dos predicados que se conectan por la palabra clave AND. El primer predicado (DERECHOSDEAUTOR > 1989) especifica que todas las filas incluidas en los resultados de la consulta deben contener un valor mayor que 1989 en la columna DERECHOSDEAUTOR. El segundo predicado (DERECHOSDEAUTOR < 2000) especifica que todas las filas incluidas en los resultados de la consulta deben contener un valor menor que 2000 en la columna DERECHOSDEAUTOR.

Conforme se evalúan las filas, cada predicado se evalúa sobre una base individual para determinar si la fila cumple con la condición definida por ese predicado. Regresando al último ejemplo, el primer predicado establece la condición de que los valores deben ser mayores que 1989. Si el valor DERECHOSDEAUTOR para una fila en particular es mayor que 1989, la condición se cumple y el predicado evalúa como verdadero. Si el valor no es mayor que 1989, el predicado evalúa como falso. Si SQL no puede determinar si el valor cumple o no la condición (como sería el caso si el valor es nulo), el predicado evalúa como desconocido.

Cada predicado evalúa como verdadero, falso o desconocido. Si se incluye más de un predicado en la cláusula WHERE, éstos se unen por la palabra clave OR o por la palabra clave AND. Si se utiliza OR, entonces al menos uno de los predicados de cualquier lado de OR debe evaluar como verdadero para que la fila pase el filtro, y por lo tanto aparece en los resultados de la consulta. Si se utiliza AND, entonces los predicados de cualquier lado deben evaluar como verdaderos para que la fila pase el filtro. Por ejemplo, la cláusula WHERE en el último ejemplo incluye dos predicados que se conectan por la palabra clave AND. Esto significa que el primer predicado debe evaluar como verdadero y el segundo predicado debe evaluar como verdadero. Si se utiliza OR en lugar de AND, entonces sólo uno de los predicados debe evaluarse como verdadero, que carece un poco de sentido en este caso, ya que todos los valores, excepto los valores nulos, están por encima de 1989 o por debajo de 2000.

Por último, la cláusula WHERE en su conjunto debe evaluar como verdadero a fin de que una fila se incluya en los resultados de una consulta. Si la cláusula WHERE incluye más de un predicado, SQL sigue las directrices específicas para la forma en que una instrucción en conjunto se evalúa. Empecemos viendo la palabra clave OR. La tabla 7-1 enumera la evaluación de una condición de búsqueda si la palabra clave OR se utiliza para separar dos predicados. Para usar la tabla, em-

	Verdadero	Falso	Desconocido
Verdadero	Verdadero	Verdadero	Verdadero
Falso	Verdadero	Falso	Desconocido
Desconocido	Verdadero	Desconocido	Desconocido

Tabla 7-1 Evaluación de los predicados conectados por OR.

pareje una condición de la columna de la izquierda a una condición en la fila superior. El resultado (donde una fila y una columna se intersectan para formar una celda) muestra cómo la condición de búsqueda se evalúa sobre la base de cómo se evalúa cada predicado.

Como muestra la tabla, si ambos predicados evalúan como verdadero, entonces la condición de búsqueda evalúa como verdadero. Si ambos son falsos, entonces la condición de búsqueda evalúa como falso. Se proporciona una condición para cada posible encuentro. Por ejemplo, suponga que la instrucción `SELECT` incluye la siguiente cláusula `WHERE`:

```
WHERE TIPO_INTERPRETE = 'Folk' OR TIPO_INTERPRETE = 'Jazz'
```

Ahora suponga que el primer predicado en este ejemplo (`TIPO_INTERPRETE = 'Folk'`) evalúa como verdadero y el segundo predicado (`TIPO_INTERPRETE = 'Jazz'`) evalúa como falso. Esto significa que la fila evaluada contiene el valor `Folk` en la columna `TIPO_INTERPRETE`, pero no contiene el valor `Jazz` en esa columna. Ahora remítase a la tabla 7-1. Si selecciona Verdadero de la primera columna, selecciona Falso de la fila superior, y luego empareja estos dos valores (buscando dónde se intersectan), se puede observar que la condición de búsqueda evalúa como verdadero; por lo tanto, la fila se incluye en los resultados de la consulta.

Se puede hacer lo mismo con la palabra clave `AND` como se hizo con la palabra clave `OR`.

De nuevo, sólo tiene que coincidir cómo se evalúa cada predicado para determinar si la condición de búsqueda se evaluará como verdadera, falsa o desconocida. Recuerde que la condición de búsqueda debe evaluar como verdadero para que la fila se incluya en los resultados de la consulta. Como se puede observar, la palabra clave `AND` es mucho menos indulgente que la palabra clave `OR`. La única manera para que la condición de búsqueda evalúe como verdadero es que ambos predicados evalúen como verdaderos.

	Verdadero	Falso	Desconocido
Verdadero	Verdadero	Falso	Desconocido
Falso	Falso	Falso	Falso
Desconocido	Desconocido	Falso	Desconocido

Tabla 7-2 Evaluación de los predicados conectados por AND.

NOTA

La comparación de los operadores y predicados en general se discute con mayor detalle en el capítulo 9.

Si una condición de búsqueda incluye más de dos predicados, los predicados se evalúan de acuerdo con un orden escogido por el RDBMS, a menos que se utilicen paréntesis para separar las combinaciones de predicados. Mientras que el estándar SQL no especifica el orden en que múltiples predicados se evalúan, la mayoría de los productos RDBMS evalúan AND antes que OR. Por ejemplo, se tiene la instrucción SELECT que incluye la siguiente cláusula WHERE:

```
WHERE EN_EXISTENCIA = 6 OR EN_EXISTENCIA = 27 AND ID_DISQUERA = 833 OR  
ID_DISQUERA = 829
```

Observe que hay cuatro predicados en esta cláusula y no hay paréntesis. Asumiendo que se evalúa AND antes que OR, la anterior cláusula WHERE se evaluaría como si se escribiera de esta manera:

```
WHERE EN_EXISTENCIA = 6 OR (EN_EXISTENCIA = 27 AND ID_DISQUERA = 833) OR  
ID_DISQUERA = 829
```

Con el fin de evaluar como verdadero, una fila debe contener uno de los siguientes valores o conjunto de valores:

- EN_EXISTENCIA valor de 6
- EN_EXISTENCIA valor de 27 y ID_DISQUERA valor de 833
- ID_DISQUERA valor de 829

Cuando se incluyen ambas palabras clave AND y OR en la misma cláusula WHERE, *siempre* es una buena idea incluir paréntesis para asegurar que se reciba el filtrado intentado, teniendo en cuenta que los predicados entre paréntesis se evalúan siempre primero. Si los RDBMS hacen otras suposiciones, o si los paréntesis se utilizan alrededor de otros conjuntos de predicados, los resultados serán diferentes de lo que se ha visto. Por ejemplo, suponga que utiliza los paréntesis de la siguiente manera:

```
WHERE (EN_EXISTENCIA = 6 O EN_EXISTENCIA = 27) AND (ID_DISQUERA = 833 O  
ID_DISQUERA = 829)
```

Los predicados primero se evalúan dentro del contexto de los paréntesis y luego se comparan con otros predicados en consecuencia. En este caso, una fila debe contener uno de los dos valores EN_EXISTENCIA y la fila debe contener uno de los dos valores ID_DISQUERA. Como resultado, una fila debe contener uno de los siguientes conjuntos de valores para evaluar como verdadero:

- Valor de EN_EXISTENCIA de 6 y valor de ID_DISQUERA de 833
- Valor de EN_EXISTENCIA de 6 y valor de ID_DISQUERA de 829
- Valor de EN_EXISTENCIA de 27 y valor de ID_DISQUERA de 833
- Valor de EN_EXISTENCIA de 27 y valor de ID_DISQUERA de 829

NOTA

SQL incluye tres operadores que se pueden utilizar si una condición de búsqueda se vuelve demasiado complicada. Estos operadores son IS TRUE, IS FALSE e IS UNKNOWN. Por ejemplo, se puede especificar la siguiente condición de búsqueda: (PRIMER_NOMBRE = 'Joni' AND APELLIDO = 'Mitchell') IS TRUE. Esto significa que el valor PRIMER_NOMBRE de la fila devuelta debe ser Joni y el valor APELLIDO debe ser Mitchell. En otras palabras, deben evaluar como verdadero. Si se especifica IS FALSE en esta situación, el par de predicado habría resultado como falso, lo que significa que al menos uno de los dos predicados es falso (podría no ser Joni o podría no ser Mitchell).

Otra palabra clave que puede resultar útil es la palabra clave NOT, que puede utilizarse sola o junto con las palabras clave AND y OR para especificar el inverso de un predicado. Por ejemplo, la instrucción SELECT puede incluir la siguiente cláusula WHERE:

```
WHERE NOMBRE_INTERPRETE = 'Joni Mitchell' OR NOT NOMBRE_INTERPRETE =
'Kitaro'
```

En este caso, el valor NOMBRE_INTERPRETE puede ser Joni Mitchell o puede ser cualquier valor distinto de Kitaro. Por supuesto, Joni Mitchell no es igual que Kitaro, de modo que el predicado es redundante y se obtiene el mismo resultado si se elimina. Además, se obtendría el mismo resultado si se utiliza el operador de comparación mayor o menor que (<>), por lo que la cláusula completa WHERE se puede reescribir de manera más simple como:

```
WHERE NOMBRE_INTERPRETE <> 'Kitaro'
```

Defina la cláusula WHERE

Ahora que tiene una visión general de cómo definir la cláusula WHERE, pongamos juntas las cláusulas SELECT y FROM y veamos algunos ejemplos. Los ejemplos que veremos se basan en la tabla INVENTARIO, mostrada en la figura 7-3. La tabla INVENTARIO contiene cinco columnas, algunas de las cuales utilizaremos para definir nuestras condiciones de búsqueda.

El primer ejemplo que veremos incluye la cláusula WHERE, que define qué filas se pueden devolver basadas en los valores EN_EXISTENCIA:

```
SELECT * FROM INVENTARIO
WHERE EN_EXISTENCIA < 20;
```

Si ejecuta esta instrucción, los resultados serán similares a los siguientes:

ID_DISCO_COMPACTO	TITULO_CD	DERECHOSDEAUTOR	PRECIO_ MENUDEO	EN_EXISTENCIA
99301	Famous Blue Raincoat	1991	16.99	6
99303	Court and Spark	1974	14.99	18
99304	Past Light	1983	15.99	2
99305	Kojiki	1990	15.99	5
99306	That Christmas Feeling	1993	10.99	3

ID_DISCO_COMPACTO: INT	TITULO_CD: VARCHAR(60)	DERECHOSDEAUTOR: INT	PRECIO_MENUDEO: NUMERICO(5,2)	EN_EXISTENCIA: INT
99301	Famous Blue Raincoat	1991	16.99	6
99302	Blue	1971	14.99	26
99303	Court and Spark	1974	14.99	18
99304	Past Light	1983	15.99	2
99305	Kojiki	1990	15.99	5
99306	That Christmas Feeling	1993	10.99	3
99307	Patsy Cline: 12 Greatest Hits	1988	16.99	25

Figura 7-3 Tabla INVENTARIO que contiene datos relacionados con el CD.

Como se puede observar, todas menos dos filas se incluyen en los resultados de la consulta. Las filas no incluidas contienen valores EN_EXISTENCIA mayores que 20. En otras palabras, esas dos filas evalúan como falso.

Ahora tomemos la misma instrucción SELECT y refinemos la cláusula WHERE aún más. En la nueva instrucción, la cláusula WHERE incluye dos predicados que se conectan por la palabra clave AND, como se muestra en el siguiente ejemplo:

```
SELECT * FROM INVENTARIO
WHERE EN_EXISTENCIA < 20 AND PRECIO_MENUDEO < 15.00;
```

Cuando se ejecuta esta instrucción, se reciben los siguientes resultados:

ID_DISCO_COMPACTO	TITULO_CD	DERECHOSDEAUTOR	PRECIO_MENUDEO	EN_EXISTENCIA
-----	-----	-----	-----	-----
99303	Court and Spark	1974	14.99	18
99306	That Christmas Feeling	1993	10.99	3

Observe que sólo dos filas cumplen la condición de búsqueda. En otras palabras, sólo estas dos filas tienen un valor EN_EXISTENCIA menor que 20 y un valor PRECIO_MENUDEO menor que 15.00. Debido a que se utiliza la palabra clave AND, ambos predicados deben evaluar como verdadero, que se aplica para estas dos filas.

Ahora hagamos una pequeña modificación a la instrucción SELECT. En la cláusula WHERE, se cambió la palabra clave AND por las palabras clave AND NOT, como se muestra en el siguiente ejemplo:

```
SELECT * FROM INVENTARIO
WHERE EN_EXISTENCIA < 20 AND NOT PRECIO_MENUDEO < 15.00;
```

La palabra clave NOT cambia los resultados de la consulta. Como se puede observar, tres filas se devuelven:

ID_DISCO_COMPACTO	TITULO_CD	DERECHOSDEAUTOR	PRECIO_ MENUDEO	EN_EXISTENCIA
99301	Famous Blue Raincoat	1991	16.99	6
99304	Past Light	1983	15.99	2
99305	Kojiki	1990	15.99	5

Cada una de las filas devueltas contiene un valor EN_EXISTENCIA menor que 20 y un valor PRECIO_MENUDEO que *no* es menor que 15.00, o 15.00 o mayor.

Como siguiente paso veamos la misma instrucción SELECT, sólo que en esta ocasión los dos predicados se conectan por la palabra clave OR, como se muestra en el siguiente ejemplo:

```
SELECT * FROM INVENTARIO
WHERE EN_EXISTENCIA < 20 OR PRECIO_MENUDEO < 15.00;
```

Los resultados de la consulta para esta instrucción incluyen muchas más filas que cuando se utilizó la palabra clave AND. Por su naturaleza, la palabra clave OR permite que existan mayores oportunidades para que la búsqueda de la cláusula evalúe como verdadera. Como se puede observar, ahora se devuelven seis filas:

ID_DISCO_COMPACTO	TITULO_CD	DERECHOSDEAUTOR	PRECIO_ MENUDEO	EN_EXISTENCIA
99301	Famous Blue Raincoat	1991	16.99	6
99302	Blue	1971	14.99	26
99303	Court and Spark	1974	14.99	18
99304	Past Light	1983	15.99	2
99305	Kojiki	1990	15.99	5
99306	That Christmas Feeling	1993	10.99	3

Cada fila en los resultados de la búsqueda contiene un valor EN_EXISTENCIA menor que 20 o un valor PRECIO_MENUDEO menor que 15.00. Debido a que la palabra clave OR se utiliza, sólo uno de los predicados necesita evaluar como verdadero, aunque es aceptable si ambos predicados evalúan como verdadero.

En el siguiente ejemplo se añade un predicado más que limita las filas devueltas a aquellas con un valor EN_EXISTENCIA mayor que 5, puesto entre paréntesis para hacerlo claro:

```
SELECT * FROM INVENTARIO
WHERE (EN_EXISTENCIA < 20 AND EN_EXISTENCIA > 5) OR PRECIO_MENUDEO <
15.00;
```

Para devolver una fila, el valor EN_EXISTENCIA debe estar entre el rango de 5 y 20 o el valor PRECIO_MENUDEO debe ser menor que 15.00. Los resultados de la consulta de esta instrucción SELECT serían como sigue:

ID_DISCO_COMPACTO	TITULO_CD	DERECHOSDEAUTOR	PRECIO_ MENUDEO	EN_EXISTENCIA
99301	Famous Blue Raincoat	1991	16.99	6
99302	Blue	1971	14.99	26
99303	Court and Spark	1974	14.99	18
99306	That Christmas Feeling	1993	10.99	3

Ahora hagamos un cambio más a la cláusula WHERE. Suponga que desea que el valor EN_EXISTENCIA sea menor que 20 y mayor que 5 o el valor EN_EXISTENCIA sea menor que 20 y el valor PRECIO_MENUDEO sea menor que 15. Una forma de hacer esto es colocando paréntesis en los últimos dos predicados:

```
SELECT * FROM INVENTARIO
  WHERE EN_EXISTENCIA < 20 AND (EN_EXISTENCIA > 5 OR PRECIO_MENUDEO <
15.00) ;
```

Los resultados recibidos en esta ocasión son algo diferentes ya que la fila Blue ya no evalúa como verdadero:

ID_DISCO_COMPACTO	TITULO_CD	DERECHOSDEAUTOR	PRECIO_MENUDEO	EN_EXISTENCIA
99301	Famous Blue Raincoat	1991	16.99	6
99303	Court and Spark	1974	14.99	18
99306	That Christmas Feeling	1993	10.99	3

Con la combinación de predicados se pueden crear muchas condiciones de búsqueda que permitan devolver exactamente los datos necesarios. La clave para escribir condiciones de búsqueda eficaces es una comprensión profunda de los predicados y los operadores utilizados para formar esos predicados. El capítulo 9 muestra muchos de los operadores que se pueden utilizar y los tipos de predicados que se pueden crear. Con esa información se pueden crear condiciones eficaces y concisas de búsqueda.

Utilice la cláusula GROUP BY para agrupar los resultados de una consulta

La siguiente cláusula en la instrucción SELECT es la cláusula GROUP BY. La cláusula GROUP BY tiene una función muy diferente de la cláusula WHERE. Como su nombre lo indica, la cláusula GROUP BY se utiliza para agrupar tipos de información con el fin de resumir datos relacionados. La cláusula GROUP BY se puede incluir en la instrucción SELECT aun si la cláusula WHERE se utiliza o no.

Como se vio en la sección “Utilice la instrucción SELECT para la recuperación de datos”, la sintaxis para la cláusula GROUP BY, como aparece en la sintaxis de la instrucción SELECT, se ve como se muestra a continuación:

```
[ GROUP BY <especificaciones de grupo> ]
```

Sin embargo, el marcador de posición <especificaciones de grupo> se puede dividir en elementos más pequeños:

```
<nombre de columna>[ {,<nombre de columna> }... ]
[ { ROLLUP | CUBE } ( <nombre de columna> [ { , <nombre de columna> }... ] )
```

En realidad, la sintaxis <especificaciones de grupo>, como algunas de las otras sintaxis en este libro, es incluso más compleja de la que se presenta aquí; sin embargo, para propósitos de este capítulo, esta sintaxis proveerá los detalles necesarios para usar la cláusula GROUP BY de manera efectiva.

Ahora veamos la sintaxis. La primera línea se explica por sí sola. Se especifican uno o más nombres de columna que contengan valores que se agrupan juntos. Esto normalmente aplica a columnas que representan algunos tipos de categorías cuyos valores se repiten dentro de la tabla. Por ejemplo, la base de datos puede incluir una tabla que enumere a los empleados en la organización. Suponga que la tabla incluye el título del puesto para cada empleado. Posiblemente encuentre que desea agrupar la información en la tabla por el título del puesto, con una fila en el conjunto de resultados para cada valor del título del puesto, quizá para determinar ciertas cosas como el salario promedio de cada puesto o el número de empleados que tiene cada título del puesto. Si se necesita especificar más de un nombre de columna, asegúrese de separarlos con una coma seguida de cada nombre (excepto el último).

En relación con la sintaxis, se puede especificar la segunda línea en lugar de la primera. En este caso, se puede usar la palabra clave `ROLLUP` o `CUBE`, junto con el listado de los nombres de columna, entre paréntesis. De nuevo, asegúrese de separar los nombres de columna con comas. Con respecto a las palabras clave `ROLLUP` y `CUBE`, la mejor manera de entender estos operadores es mediante el uso de ejemplos. De hecho, la mejor forma de entender totalmente la cláusula `GROUP BY` es a través de ejemplos. Sin embargo, antes de entrar en esto, echemos un vistazo a la tabla en la cual se basan los ejemplos. La figura 7-4 muestra la tabla `EXISTENCIA_DISCO_COMPACTO`, que contiene una lista de CD, si son interpretados o instrumentales, el precio, y cuántos de cada título hay actualmente en existencia.

Ahora podemos empezar con los ejemplos. En el primero que veremos se usará la cláusula `GROUP BY` para agrupar filas basadas en la columna `CATEGORIA` de la tabla `EXISTENCIA_DISCO_COMPACTO`, como se muestra en la siguiente instrucción `SELECT`:

```
SELECT CATEGORIA, SUM(A_LA_MANO) AS TOTAL_A_LA_MANO
FROM EXISTENCIA_DISCO_COMPACTO
GROUP BY CATEGORIA;
```

Primero echemos un vistazo a la cláusula `GROUP BY`, que especifica qué filas se deben agrupar basadas en la columna `CATEGORIA`. Si observa la figura 7-4, verá que la columna contiene sólo dos valores: Vocal o Instrumental. Como resultado, la instrucción `SELECT` sólo devuelve dos columnas, una para Instrumental y otra para Vocal:

CATEGORIA	TOTAL_A_LA_MANO
Instrumental	78
Vocal	217

Ahora veamos la cláusula `SELECT` en el ejemplo anterior de la instrucción `SELECT`. Observe que la lista seleccionada incluye la función `SUM`, que añade información a la columna `A_LA_MANO`. La columna resultante se llama `TOTAL_A_LA_MANO`. La otra columna incluida en la lista seleccionada es la columna `CATEGORIA`. La lista seleccionada puede incluir sólo esas columnas que se especifican en la cláusula `GROUP BY` o que de alguna manera se pueden resumir.

Lo que hace esta instrucción, entonces, es añadir los valores totales `A_LA_MANO` para cada valor en la columna `CATEGORIA`. En este caso, hay un total de 217 CD en existencia que entran en la categoría de interpretados, y 78 en existencia que entran en la categoría de instrumentales. Si hubiera otra categoría, entonces aparecería otra fila para esa también.

DISCO_COMPACTO: VARCHAR(60)	CATEGORIA: VARCHAR(15)	PRECIO: NUMERICO(5,2)	A_LA_MANO: INT
Famous Blue Raincoat	Vocal	16.99	13
Blue	Vocal	14.99	42
Court and Spark	Vocal	14.99	22
Past Light	Instrumental	15.99	17
Kojiki	Instrumental	15.99	6
That Christmas Feeling	Vocal	14.99	8
Patsy Cline: 12 Greatest Hits	Vocal	16.99	32
Carreras Domingo Pavarotti in Concert	Vocal	15.99	27
After the Rain: The Soft Sounds of Erik Satie	Instrumental	16.99	21
Out of Africa	Instrumental	16.99	29
Leonard Cohen The Best of	Vocal	15.99	12
Fundamental	Vocal	15.99	34
Blues on the Bayou	Vocal	14.99	27
Orlando	Instrumental	14.99	5

Figura 7-4 Información del CD en la tabla EXISTENCIA_DISCO_COMPACTO.

Pregunta al experto

P: ¿Existen consideraciones de ejecución con respecto al uso de la cláusula GROUP BY?

R: Sí, el uso de GROUP BY puede causar problemas de ejecución, ya que los RDBMS por lo general deben realizar un ordenamiento adecuado con el fin de agrupar los grupos de filas, y ordenar un gran número de filas (decenas de miles o más) puede consumir considerables recursos. Pero también esté consciente que la cláusula ORDER BY (presentada más adelante en este capítulo) y la palabra clave DISTINCT también suelen requerir ordenaciones, y por lo tanto tienen consideraciones de ejecución similares. Esto no significa que deba tener miedo de utilizarlos, sino que debe esforzarse por aprender el impacto de ejecución de las instrucciones de SQL para ganar experiencia, y de esa manera estar más capacitado al escribir instrucciones que tengan una mejor ejecución en la implementación del proveedor. Por ejemplo, en Oracle, la cláusula GROUP BY que enumera todas las columnas en la cláusula SELECT es más eficaz que utilizar la palabra clave DISTINCT, y sin embargo los resultados de la consulta de los dos métodos son idénticos.

Como se dijo anteriormente, se puede utilizar la cláusula `WHERE` en la instrucción `SELECT` que incluya la cláusula `GROUP BY`. Por ejemplo, suponga que desea ver los totales sólo de los CD que se vendieron por menos de \$16.00. Para hacer esto, simplemente modifique la instrucción `SELECT` como se muestra a continuación:

```
SELECT CATEGORIA, SUM(A_LA_MANO) AS TOTAL_A_LA_MANO
FROM EXISTENCIA_DISCO_COMPACTO
WHERE PRICE < 16.00
GROUP BY CATEGORIA;
```

Los resultados de la consulta de esta instrucción serían ligeramente diferentes si no se incluyera la cláusula `WHERE`:

CATEGORIA	TOTAL_A_LA_MANO
-----	-----
Instrumental	28
Vocal	172

Observe que como los CD que se venden por \$16.00 o más se excluyen, los resultados ahora muestran sólo 28 CD instrumental y 172 CD vocal.

En los dos ejemplos anteriores, la cláusula `GROUP BY` especifica sólo una columna. Sin embargo, se pueden especificar columnas adicionales como sea necesario. Esto permite crear subgrupos que agrupen datos en el ámbito de los grupos principales. Por ejemplo, suponga que desea agrupar datos no sólo por los valores en la columna `CATEGORIA`, sino también de acuerdo con los valores en la columna `PRECIO`. Para hacer esto, se debe incluir la columna `PRECIO` en la lista de selección, así como en la cláusula `GROUP BY`, como se muestra en la siguiente instrucción `SELECT`:

```
SELECT CATEGORIA, PRECIO, SUM(A_LA_MANO) AS TOTAL_A_LA_MANO
FROM EXISTENCIA_DISCO_COMPACTO
GROUP BY CATEGORIA, PRECIO;
```

Ahora los resultados de la consulta incluyen seis filas, en lugar de dos:

CATEGORIA	PRECIO	TOTAL_A_LA_MANO
-----	-----	-----
Instrumental	14.99	5
Vocal	14.99	99
Instrumental	15.99	23
Vocal	15.99	73
Instrumental	16.99	50
Vocal	16.99	45

Observe que para cada valor en `CATEGORIA`, hay tres filas, una para cada uno de los valores `PRECIO`. Por ejemplo, en el grupo Vocal, hay 99 CD a 14.99, 73 CD a 15.99 y 45 CD a 16.99. El número de filas depende de cuántos valores diferentes existan en las columnas especificadas en la cláusula `GROUP BY`. En este ejemplo, hay dos valores diferentes en la columna `CATEGORIA` y tres valores diferentes en la columna `PRECIO`, lo que significa que seis (dos veces tres) filas se devuelven.

NOTA

El orden en el que los resultados de la consulta se devuelven puede variar de una aplicación a otra. Por ejemplo, algunos productos pueden devolver todas las filas de instrumental juntas, seguidas por todas las filas de los interpretados. Sin embargo, independientemente de cómo aparece la información en la interfaz de usuario, los resultados finales deben ser los mismos. También la mayoría de las aplicaciones permiten añadir la cláusula ORDER BY (que se analiza después en el capítulo) para ordenar las filas resumidas.

Ahora echemos un vistazo a los operadores ROLLUP y CUBE. Ambos operadores son similares en funcionamiento, ya que devuelven datos adicionales en los resultados de la consulta cuando se añade la cláusula GROUP BY. La principal diferencia entre los dos es que el operador CUBE devuelve más información que el operador ROLLUP. Empecemos con el ejemplo del operador ROLLUP para que se pueda demostrar la diferencia.

En la siguiente instrucción SELECT, en la cláusula GROUP BY aplica el operador ROLLUP a las columnas CATEGORIA y PRECIO:

```
SELECT CATEGORIA, PRECIO, SUM(A_LA_MANO) AS TOTAL_A_LA_MANO
FROM EXISTENCIA_DISCO_COMPACTO
GROUP BY ROLLUP (CATEGORIA, PRECIO);
```

NOTA

Las implementaciones pueden variar con respecto a cómo soportan a los operadores ROLLUP y CUBE. Por ejemplo, en SQL Server se debe añadir WITH ROLLUP o WITH CUBE al final de la cláusula GROUP BY, en lugar de definir la cláusula en la forma que la especifica el estándar SQL:2006. Asegúrese de verificar la documentación del producto para determinar cómo se respaldan estos operadores.

Ahora cuando se ejecute la instrucción SELECT, los resultados de la consulta incluyen una fila adicional para cada valor en la columna CATEGORIA, además de una fila con el gran total al final:

CATEGORIA	PRECIO	TOTAL_A_LA_MANO
-----	-----	-----
Instrumental	14.99	5
Instrumental	15.99	23
Instrumental	16.99	50
Instrumental	NULL	78
Vocal	14.99	99
Vocal	15.99	73
Vocal	16.99	45
Vocal	NULL	217
NULL	NULL	295

Las dos filas adicionales de CATEGORIA proporcionan los totales para cada valor en la columna CATEGORIA. En el ejemplo anterior, el grupo Instrumental incluye un total de 78 CD, y el grupo Vocal incluye un total de 217 CD. Observe que la columna PRECIO incluye un valor nulo para estas filas en particular. Un valor no se puede calcular para esta columna, ya que los tres subgrupos

(de la columna PRECIO) se representan aquí. La última fila (la fila con NULL para ambas columnas CATEGORIA y PRECIO) contiene el gran total de todos los CD incluidos por la consulta (ambas categorías de grupos y los tres subgrupos de precios).

El operador CUBE devuelve los mismos datos que el operador ROLLUP y algunos otros. Observe que en la siguiente instrucción SELECT, simplemente se reemplazó la palabra clave CUBE por ROLLUP:

```
SELECT CATEGORIA, PRECIO, SUM(A_LA_MANO) AS TOTAL_A_LA_MANO
FROM EXISTENCIA_DISCO_COMPACTO
GROUP BY CUBE (CATEGORIA, PRECIO);
```

Esta instrucción devuelve los siguientes resultados de la consulta:

CATEGORIA	PRECIO	TOTAL_A_LA_MANO
-----	-----	-----
Instrumental	14.99	5
Instrumental	15.99	23
Instrumental	16.99	50
Instrumental	NULL	78
Vocal	14.99	99
Vocal	15.99	73
Vocal	16.99	45
Vocal	NULL	217
NULL	NULL	295
NULL	14.99	104
NULL	15.99	96
NULL	16.99	95

Se puede observar que tres filas adicionales se añaden a los resultados de la consulta, una fila para cada valor diferente en la columna PRECIO. A diferencia del operador ROLLUP, el operador CUBE resume los valores para cada subgrupo. También observe que un valor nulo se muestra para la columna CATEGORIA para esas filas. Esto es debido a que ambos valores (vocal e instrumental) se incluyen en el resumen de cada subgrupo.

Como se puede observar, la cláusula GROUP BY puede ser una herramienta muy útil cuando se trata de resumir datos, particularmente cuando se hace uso de las múltiples funciones disponibles en SQL, tales como SUM y AVG. En el capítulo 10 se analizan éstas y muchas otras funciones que se pueden utilizar para hacer la instrucción SELECT más robusta y aplicable a sus necesidades.

Utilice la cláusula HAVING para especificar un grupo de condiciones de búsqueda

La cláusula HAVING es similar a la cláusula WHERE ya que define una condición de búsqueda. Sin embargo, a diferencia de la cláusula WHERE, la cláusula HAVING se refiere a grupos, no a filas individuales:

- Si se especifica la cláusula GROUP BY, la cláusula HAVING se aplica a los grupos creados por la cláusula GROUP BY.

- Si se especifica la cláusula WHERE y no se especifica la cláusula GROUP BY, la cláusula HAVING se aplica a la salida de la cláusula WHERE y se trata esa salida como un grupo.
- Si no se especifica la cláusula WHERE ni la cláusula GROUP BY, la cláusula HAVING se aplica a la salida de la cláusula FROM y se trata esa salida como un grupo.

La mejor forma de entender la cláusula HAVING es recordando que las cláusulas en la instrucción SELECT se procesan con un orden definido. La cláusula WHERE sólo puede recibir la entrada de la cláusula FROM, pero la cláusula HAVING puede recibir la entrada de las cláusulas GROUP BY, WHERE o FROM. Ésta es una sutil pero importante distinción, y la mejor manera de ilustrarlo es viendo un par de ejemplos.

En el primer ejemplo, que se basa en la tabla EXISTENCIA_DISCO_COMPACTO en la figura 7-4, se utiliza la cláusula WHERE para especificar que los resultados de la consulta deben incluir sólo filas cuyo valor A_LA_MANO es menor que 20, como se muestra en la siguiente instrucción SELECT:

```
SELECT CATEGORIA, AVG(PRECIO) AS PROM_PRECIO
FROM EXISTENCIA_DISCO_COMPACTO
WHERE A_LA_MANO < 20
GROUP BY CATEGORIA;
```

La instrucción devuelve dos columnas: CATEGORIA y PROM_PRECIO, que es el promedio de todos los precios para cada categoría. Los promedios incluyen sólo las columnas donde los valores A_LA_MANO son menores que 20. Si se ejecuta esta instrucción, los resultados serían similares a los siguientes:

CATEGORIA	PROM_PRECIO
Instrumental	15.656666
Vocal	15.990000

Como era de esperarse, el resultado de la consulta devuelve dos filas (una para el grupo instrumental y una para el grupo vocal).

Si se fuera a utilizar la cláusula HAVING, en lugar de la cláusula WHERE, para limitar los valores a menos de 20, podría utilizar la siguiente instrucción SELECT:

```
SELECT CATEGORIA, AVG(PRECIO) AS PROM_PRECIO
FROM EXISTENCIA_DISCO_COMPACTO
GROUP BY CATEGORIA
HAVING A_LA_MANO < 20;
```

Sin embargo, si se trata de ejecutar esta instrucción, se recibiría un error, ya que no se aplican los valores individuales A_LA_MANO a los grupos. Para que una columna se incluya en la cláusula HAVING, debe ser un grupo de columnas o deben estar resumidas de alguna manera.

Ahora veamos otro ejemplo que utilice la cláusula HAVING. En este caso, la cláusula incluye una columna resumida:

```
SELECT PRECIO, CATEGORIA, SUM(A_LA_MANO) AS TOTAL_A_LA_MANO
FROM EXISTENCIA_DISCO_COMPACTO
GROUP BY PRECIO, CATEGORIA
HAVING SUM(A_LA_MANO) > 10;
```

La cláusula HAVING en esta instrucción funciona porque los valores A_LA_MANO se suman, lo que significa que pueden trabajar dentro de la estructura del grupo. Los resultados de la consulta serían como se muestra a continuación:

PRECIO	CATEGORIA	TOTAL_A_LA_MANO
-----	-----	-----
15.99	Instrumental	23
16.99	Instrumental	50
14.99	Vocal	99
15.99	Vocal	73
16.99	Vocal	45

La cláusula HAVING se aplica a los resultados después de haberse agrupado (en la cláusula GROUP BY). Para cada grupo, los valores A_LA_MANO se añaden juntos, pero sólo se incluyen los grupos con los valores TOTAL_A_LA_MANO de más de 10. Si no se incluye la cláusula HAVING, los resultados de la consulta se incluirían en una fila adicional para el grupo de 14.99/Instrumental.

En su mayor parte, probablemente encuentre que utilizará la cláusula HAVING en conjunto con la cláusula GROUP BY. Mediante el uso de estos dos se pueden agrupar datos relevantes y filtrar los datos para refinar su búsqueda aún más. La cláusula HAVING también tiene la ventaja de permitir el uso de funciones establecidas tales como AVG o SUM, que no se pueden utilizar en la cláusula WHERE a menos que se coloquen dentro de una subconsulta. Los puntos importantes que deben tenerse en cuenta con la cláusula HAVING es que es la última cláusula en la expresión de tabla que debe aplicarse, y que se trata de datos agrupados en lugar de filas individuales.

Utilice la cláusula ORDER BY para ordenar los resultados de una consulta

La cláusula ORDER BY, cuando se usa en la instrucción SELECT, es la última cláusula procesada. La cláusula ORDER BY toma la salida de la cláusula SELECT y ordena los resultados de la consulta de acuerdo con las especificaciones dentro de la cláusula ORDER BY. La cláusula no agrupa las filas, como se agrupan por la cláusula GROUP BY, ni filtra las filas, como se filtran por la cláusula WHERE o la cláusula HAVING. Sin embargo, se puede especificar si las filas se organizan en un orden ascendente (utilizando la palabra clave ASC) o en orden descendente (usando la palabra clave DESC).

Para utilizar la cláusula ORDER BY, simplemente se especifica una o más columnas y las palabras clave opcionales ASC o DESC (una por columna). Si no se especifica la palabra clave, se toma ASC. Las filas se organizan de acuerdo con la columna que especifique. Si se define más de una columna en la cláusula ORDER BY, las filas se organizan en el orden en que las columnas se especifican.

Echemos un vistazo a algunos ejemplos para aclarar cómo funciona la cláusula ORDER BY. (Los ejemplos se basan en la tabla EXISTENCIA_DISCO_COMPACTO en la figura 7-4.) En el primer ejemplo se ordenan las filas basadas en la columna PRECIO:

```
SELECT * FROM EXISTENCIA_DISCO_COMPACTO
WHERE PRECIO < 16.00
ORDER BY PRECIO;
```

Observe que la columna PRECIO se especifica en la cláusula ORDER BY. También observe que si no se especifica la palabra clave ASC ni la palabra clave DESC, entonces se asume la palabra clave ASC. Si se ejecuta esta consulta, se recibirán resultados similares a los siguientes:

DISCO_COMPACTO	CATEGORIA	PRECIO	A_LA_MANO
-----	-----	-----	-----
Blue	Vocal	14.99	42
Court and Spark	Vocal	14.99	22
That Christmas Feeling	Vocal	14.99	8
Blues on the Bayou	Vocal	14.99	27
Orlando	Instrumental	14.99	5
Carreras Domingo Pavarotti in Concert	Vocal	15.99	27
Leonard Cohen The Best Of	Vocal	15.99	12
Fundamental	Vocal	15.99	34
Past Light	Instrumental	15.99	17
Kojiki	Instrumental	15.99	6

Las filas se enumeran de acuerdo con la columna PRECIO. Los valores en la columna PRECIO aparecen en orden ascendente (del precio más bajo al precio más alto). Ya que la cláusula WHERE se especifica, ninguna fila con precios por encima de 15.99 se incluye en los resultados de la consulta. Ya que se incluye sólo la columna PRECIO en ORDER BY, también el orden de las filas que tienen el mismo precio es impredecible. Por ejemplo, las cinco filas con un PRECIO de 14.99, aparecen todas antes que aquellas con un PRECIO de 15.99, pero esas cinco filas pueden aparecer en cualquier orden.

En el siguiente ejemplo, la instrucción SELECT es casi la misma que la última instrucción, excepto que se especifica una columna adicional en la cláusula ORDER BY:

```
SELECT * FROM EXISTENCIA_DISCO_COMPACTO
WHERE PRECIO < 16.00
ORDER BY PRECIO, A_LA_MANO DESC;
```

En este caso, a la columna A_LA_MANO le sigue la palabra clave DESC, lo que significa que las filas se enumeran en orden descendente. Sin embargo, ya que hay dos columnas, las filas primero se ordenan por la columna PRECIO y luego por la columna A_LA_MANO. Si se ejecuta esta instrucción SELECT, se recibirán los siguientes resultados:

DISCO_COMPACTO	CATEGORIA	PRECIO	A_LA_MANO
-----	-----	-----	-----
Blue	Vocal	14.99	42
Blues on the Bayou	Vocal	14.99	27
Court and Spark	Vocal	14.99	22
That Christmas Feeling	Vocal	14.99	8
Orlando	Instrumental	14.99	5
Fundamental	Vocal	15.99	34
Carreras Domingo Pavarotti in Concert	Vocal	15.99	27
Past Light	Instrumental	15.99	17
Leonard Cohen The Best Of	Vocal	15.99	12
Kojiki	Instrumental	15.99	6

Pregunta al experto

P: ¿Cómo afecta la cláusula **ORDER BY** los resultados de la consulta en SQL incrustado y módulos de SQL?

R: Se puede utilizar la cláusula **ORDER BY** sólo en invocación directa y cuando se definen cursores. (Se analizarán los cursores en el capítulo 15.) No se puede utilizar la cláusula **ORDER BY** en otras situaciones. Esto es debido a las limitaciones en los lenguajes de aplicación (no pueden manejar un número indeterminado de filas de un resultado de la consulta). Los lenguajes de aplicación no saben qué hacer con este tipo de incertidumbre. Y ya que la cláusula **ORDER BY** aplica sólo a los resultados de la consulta de múltiples filas, la cláusula no es aplicable a los entornos que requieren que las filas se devuelvan una a la vez. Sin embargo, los cursores ofrecen una forma para que los lenguajes de aplicación hagan frente a esa incertidumbre, permitiendo que la cláusula **ORDER BY** se utilice en definiciones de cursores. Los cursores se analizan con más detalle en el capítulo 15.

Como se puede observar, las filas se enumeran según el orden de los valores de **PRECIO**, que es de forma ascendente. Además, los valores **A_LA_MANO** se enumeran con un orden descendente para cada precio. Por lo tanto, para el conjunto de valores **PRECIO** de 14.99, las filas comienzan con el valor de 42 en la columna **A_LA_MANO** y terminan con el valor de 5. Luego brinca al siguiente grupo de valores **PRECIO**: 15.99. Una vez más, el valor **A_LA_MANO** más grande para el rango de **PRECIO** de 15.99 aparece en primer lugar y la última fila contiene el valor **A_LA_MANO** más pequeño para el rango de **PRECIO** de 15.99.

Siempre que se utilice la cláusula **ORDER BY**, debe estar consciente del orden en el que enumera los nombres de columna dentro de la cláusula. En el ejemplo anterior, la columna **PRECIO** aparece en primer lugar; por lo tanto, las filas se ordenan primero por la columna **PRECIO** y luego por la columna **A_LA_MANO**. Sin embargo, puede invertir los nombres de columna, como se muestra en la siguiente instrucción **SELECT**:

```
SELECT * FROM EXISTENCIA_DISCO_COMPACTO
WHERE PRECIO < 16.00
ORDER BY A_LA_MANO, PRECIO DESC;
```

Esta vez, la columna **A_LA_MANO** aparece en primer lugar y la columna **PRECIO** aparece en segundo lugar, y a la columna **PRECIO** se le asigna la palabra clave **DESC**. Como resultado, las filas se ordenan primero por la columna **A_LA_MANO**, como se muestra en los siguientes resultados de la consulta:

DISCO_COMPACTO	CATEGORIA	PRECIO	A_LA_MANO
-----	-----	-----	-----
Orlando	Instrumental	14.99	5
Kojiki	Instrumental	15.99	6
That Christmas Feeling	Vocal	14.99	8
Leonard Cohen The Best Of	Vocal	15.99	12
Past Light	Instrumental	15.99	17
Court and Spark	Vocal	14.99	22

Carreras Domingo Pavarotti in Concert	Vocal	15.99	27
Blues on the Bayou	Vocal	14.99	27
Fundamental	Vocal	15.99	34
Blue	Vocal	14.99	42

Observe que los valores A_LA_MANO están en orden ascendente. Las filas se ordenan según el valor PRECIO. Sin embargo, ya que sólo hay dos filas que comparten los mismos valores A_LA_MANO (27), éstas son las únicas filas que afecta la columna ORDER BY con respecto a la columna PRECIO.

La cláusula ORDER BY es una herramienta conveniente para organizar los resultados de la consulta, pero recuerde que no afecta qué datos se muestran. Sólo las otras cláusulas pueden nombrar, filtrar y agrupar datos. La cláusula ORDER BY no es más que un organizador para lo que ya existe. Y de hecho, si bien no es una práctica muy buena, se pueden incluir columnas en la cláusula ORDER BY que no están en la cláusula SELECT, y de este modo clasificar las columnas que no son visibles en los resultados de la consulta.

Pruebe esto 7-1 **Consulte la base de datos inventario**

Para los ejercicios de los capítulos anteriores se creó un número continuo de tablas base capaces de almacenar datos. En este capítulo aprenderá cómo crear instrucciones SELECT que permitan consultar datos en las tablas base. Como resultado, este ejercicio se centra en la creación de instrucciones SELECT que consulte datos en las tablas que se crearon. Sin embargo, antes de que realmente pueda consultar estas tablas, los datos se deben almacenar dentro de éstas. Ya que no se cubre la inserción de datos hasta el capítulo 8, se proporcionan las instrucciones necesarias para insertar los datos en el archivo Try_This_07.txt (en inglés), que puede descargar del sitio web. El archivo contiene una serie de instrucciones INSERT que permiten llenar las tablas, junto con las instrucciones SELECT utilizadas en este ejercicio. También se pueden ver estas instrucciones en el apéndice C.

Si observa el archivo Try_This_07.txt, encontrará una serie de instrucciones INSERT que se agrupan de acuerdo con las tablas que se crearon en la base de datos INVENTARIO. Por ejemplo, el primer conjunto de instrucciones INSERT son para la tabla DISQUERAS_CD, como se muestra en las siguientes instrucciones:

```
--Inserta datos en la tabla DISQUERAS_CD
INSERT INTO DISQUERAS_CD VALUES ( 827, 'Private Music' );
INSERT INTO DISQUERAS_CD VALUES ( 828, 'Reprise Records' );
INSERT INTO DISQUERAS_CD VALUES ( 829, 'Asylum Records' );
INSERT INTO DISQUERAS_CD VALUES ( 830, 'Windham Hill Records' );
INSERT INTO DISQUERAS_CD VALUES ( 831, 'Geffen' );
INSERT INTO DISQUERAS_CD VALUES ( 832, 'MCA Records' );
INSERT INTO DISQUERAS_CD VALUES ( 833, 'Decca Record Company' );
INSERT INTO DISQUERAS_CD VALUES ( 834, 'CBS Records' );
INSERT INTO DISQUERAS_CD VALUES ( 835, 'Capitol Records' );
INSERT INTO DISQUERAS_CD VALUES ( 836, 'Sarabande Records' );
--Fin de inserción en la tabla DISQUERAS_CD
```

Tendrá que copiar estas instrucciones en la aplicación de cliente y ejecutarlas. Cada instrucción INSERT añade una fila de datos a la tabla aplicable. Por ejemplo, la primera instrucción INSERT que aparece en el código anterior añade una fila de datos a la tabla DISQUERAS_CD. Los valores que se añaden son 827 (para la columna ID_DISQUERA) y Private Music (para la columna NOMBRE_COMPANIA). De nuevo se presenta la instrucción INSERT con más detalle en el capítulo 8. Si se siente incómodo insertando estos datos antes de leer acerca de la instrucción INSERT, le sugiero que revise la información del capítulo 8 antes de trabajar en este ejercicio y luego vuelva aquí para realizar cada paso. Sin embargo, si decide hacer este ejercicio ahora, entonces simplemente necesita ejecutar cada instrucción, como se describe en los siguientes pasos.

NOTA

Como probablemente ha notado, cada bloque de instrucciones INSERT comienza y termina con una línea que empieza con guiones dobles (--). Los guiones dobles indican que la línea de texto que sigue es un comentario. La aplicación de SQL simplemente omite estas líneas. Los comentarios están sólo para proporcionar información a los programadores de SQL para que puedan entender mejor el código.

Paso a paso

1. Abra la aplicación de clientes de su RDBMS y conecte la base de datos INVENTARIO.
2. Abra el archivo Try_This_07.txt y copie las instrucciones INSERT en la aplicación de cliente. La mayoría de las aplicaciones permiten ejecutar bloques de instrucciones en lugar de tener que introducir los datos de una fila a la vez. Si la aplicación soporta la ejecución de múltiples instrucciones, ejecute las instrucciones una tabla a la vez copiando y pegando los bloques de instrucciones en la aplicación de clientes. Debe introducir los datos para cada tabla en el orden que los datos aparecen en el archivo Try_This_07.txt. Por ejemplo, debe insertar los valores en la tabla DISQUERAS_CD antes que en la tabla DISCO_COMPACTO.

Para cada instrucción INSERT que ejecute, debe recibir un mensaje que reconozca que la fila se insertó en la tabla. Después de llenar cada tabla con los datos, estará listo para continuar con el siguiente paso.

3. Ahora consulte todos los datos de la tabla ARTISTAS. Introduzca y ejecute la siguiente instrucción SQL:

```
SELECT *  
FROM ARTISTAS;
```

El resultado de su consulta debe incluir las columnas ID_ARTISTA, NOMBRE_ARTISTA y LUGAR_DE_NACIMIENTO. Debe componerse de 18 filas de datos en todas.

4. Ahora debe crear una consulta que especifique qué columnas se incluyen en los resultados de la consulta. Para la siguiente instrucción SELECT, consulte la tabla DISCO_COMPACTO, pero devuelva sólo las columnas TITULO_CD y EN_EXISTENCIA. Introduzca y ejecute la siguiente instrucción SQL:

```
SELECT TITULO_CD, EN_EXISTENCIA  
FROM DISCOS_COMPACTOS;
```

Los resultados de la consulta deben incluir sólo las dos columnas que se especifican en la instrucción SELECT. Además, la consulta debe devolver 15 filas de datos.

5. En el capítulo 5, Pruebe esto 5-1, se creó la vista CD_EN_EXISTENCIA. La vista devuelve los mismos datos que se especifican en la instrucción SELECT en el paso 4, excepto que limita los resultados de las filas con los valores EN_EXISTENCIA mayor que 10. Ahora puede consultar esa vista. Introduzca y ejecute la siguiente instrucción SQL:

```
SELECT *
FROM CD_EN_EXISTENCIA;
```

Observe que la instrucción SELECT es la misma que si fuera para una tabla base persistente. Puede incluso especificar los nombres de columna de la vista si se desea. (De hecho, debe hacerlo si consulta la vista de cualquier otra manera que a través de la invocación directa.) En la última instrucción SELECT, la consulta devolvió 15 filas, pero esta consulta devuelve sólo 12 filas, ya que los valores EN_EXISTENCIA deben estar por arriba de 10. La parte agradable sobre la vista es que ya está configurada para devolver exactamente la información deseada, sin tener que definir la cláusula WHERE.

6. Ahora consulte la tabla DISCOS_COMPACTOS, pero refine la instrucción SELECT utilizando la cláusula WHERE. Introduzca y ejecute la siguiente instrucción SQL:

```
SELECT TITULO_CD, EN_EXISTENCIA
FROM DISCOS_COMPACTOS
WHERE EN_EXISTENCIA > 10 AND EN_EXISTENCIA < 30;
```

Debido a que se añade la cláusula WHERE, los resultados de la consulta ahora incluyen sólo nueve filas, y cada fila debe contener un valor EN_EXISTENCIA entre 10 y 30.

7. La siguiente instrucción SELECT que se crea agrupa la información en la tabla DISCOS_COMPACTOS. Introduzca y ejecute la siguiente instrucción SQL:

```
SELECT ID_DISQUERA, SUM(EN_EXISTENCIA) AS TOTAL_EN_EXISTENCIA
FROM DISCOS_COMPACTOS
GROUP BY ID_DISQUERA;
```

Se devuelve una fila para cada valor diferente ID_DISQUERA, y para cada uno de estos valores se devuelve el total para los valores EN_EXISTENCIA. Hay 10 filas. Observe que en los resultados de la consulta, el nombre de la columna con los totales EN_EXISTENCIA es TOTAL_EN_EXISTENCIA. Cuando aprenda más sobre la unión de tablas, será capaz de agrupar los datos basados en consultas más complejas. La unión de tablas se analiza en el capítulo 11.

8. Ahora se añade la cláusula HAVING a la instrucción SELECT que se acaba de ejecutar. Introduzca y ejecute la siguiente instrucción SQL:

```
SELECT ID_DISQUERA, SUM(EN_EXISTENCIA) AS TOTAL_EN_EXISTENCIA
FROM DISCOS_COMPACTOS
GROUP BY ID_DISQUERA
HAVING SUM(EN_EXISTENCIA) > 10;
```

(continúa)

La cláusula HAVING limita las filas que se devuelven a aquellas cuyos valores TOTAL_EN_EXISTENCIA sean mayores que 10. Ahora sólo se devuelven ocho filas.

9. También se puede ejecutar la instrucción SELECT que ordene los datos devueltos por la consulta. Introduzca y ejecute la siguiente instrucción SQL:

```
SELECT *  
  FROM DISCOS_COMPACTOS  
 WHERE EN_EXISTENCIA > 10  
 ORDER BY TITULO_CD DESC;
```

Los resultados de la consulta deben estar organizados de acuerdo con la columna TITULO_CD, con las columnas enumeradas en orden descendente. Ya que se utiliza la cláusula WHERE, sólo 12 filas se deben devolver.

10. Cierre la aplicación clientes.

Resumen de Pruebe esto

En este ejercicio se insertaron datos en las tablas de la base de datos INVENTARIO. A continuación se crearon instrucciones SELECT que permiten consultar los datos en esas tablas. Debe sentirse libre de experimentar con las instrucciones SELECT y probar diferentes tipos de consultas. A medida que se sienta más cómodo utilizando la instrucción SELECT y aprenda técnicas más avanzadas para consultar datos, será capaz de escribir instrucciones SELECT para acceder a múltiples tablas, calcular datos y resumir información. Sin embargo, incluso las técnicas más avanzadas se basan en el fundamento básico que ha demostrado en este ejercicio. Todo lo demás se basa en esto.

Autoexamen Capítulo 7

1. ¿Cuáles cláusulas en la instrucción SELECT son parte de la expresión de la tabla?
A SELECT
B FROM
C WHERE
D ORDER BY
2. ¿En qué orden se aplican las cláusulas de la instrucción SELECT?
3. Se está escribiendo la instrucción SELECT que recupera la columna TITULO_CD y todas las filas de la tabla INVENTARIO. ¿Cuál instrucción SELECT debe utilizarse?
4. Se está escribiendo una instrucción SELECT que recupera la columna TITULO_CD y todas las filas de la tabla INVENTARIO. Desea que la columna en los resultados de la consulta sea llamada DISCO_COMPACTO. ¿Cuál instrucción SELECT deberá utilizarse?

5. ¿Cuáles cláusulas se requieren en una instrucción SELECT?
 - A SELECT
 - B FROM
 - C WHERE
 - D GROUP BY
6. ¿Cuál palabra clave se debe añadir a la cláusula SELECT para asegurarse que cada fila de los resultados de la consulta sea única?
 - A ALL
 - B ROLLUP
 - C DISTINCT
 - D CUBE
7. Se está creando una instrucción SELECT para la tabla INVENTARIO y desea asegurarse que sólo las filas con un valor PRECIO_MENUDEO menor a \$16.00 sean incluidas en los resultados de la consulta. ¿Qué cláusula WHERE deberá utilizarse?
8. Se está creando una instrucción SELECT que incluye una cláusula WHERE. La cláusula WHERE contiene dos predicados. Se desea que la condición de uno de los predicados se cumpla, pero no es necesario que ambas condiciones se cumplan. ¿Qué palabra clave deberá utilizarse para conectar los dos predicados?
9. ¿Cuál de las siguientes declaraciones puede evaluar una cláusula WHERE?
 - A Verdadero
 - B No
 - C Falso
 - D Desconocido
10. ¿Cuál cláusula permite agrupar valores en una columna específica?
 - A ROLLUP
 - B HAVING
 - C ORDER BY
 - D GROUP BY
11. ¿Cuáles dos operadores pueden utilizarse en una cláusula GROUP BY para arrojar datos de resumen adicionales en los resultados de una consulta?
 - A ROLLUP
 - B HAVING
 - C CUBE
 - D DISTINCT

- 12.** Se está escribiendo la instrucción SELECT que recupera las columnas CATEGORIA y PRECIO de la tabla EXISTENCIA_DISCO_COMPACTO. Quiere agrupar los datos primero por la columna CATEGORIA y luego por la columna PRECIO. ¿Cuál instrucción SELECT deberá utilizarse?
- 13.** Se está escribiendo la instrucción SELECT que recupera las columnas CATEGORIA y PRECIO de la tabla EXISTENCIA_DISCO_COMPACTO. Quiere agrupar los datos primero por la columna CATEGORIA y luego por la columna PRECIO. A continuación, desea filtrar cualquier grupo que tenga un valor PRECIO superior a \$15.99. ¿Cuál instrucción SELECT deberá utilizarse?
- 14.** Se crea una instrucción SELECT que incluye una cláusula SELECT, una cláusula FROM, una cláusula WHERE, una cláusula GROUP BY y una cláusula HAVING. ¿Desde cuál cláusula recibirá resultados la cláusula HAVING?
- A** SELECT
 - B** FROM
 - C** WHERE
 - D** GROUP BY
- 15.** ¿En qué aspecto la cláusula HAVING es diferente de la cláusula WHERE?
- 16.** ¿De cuál cláusula recibe resultados la cláusula ORDER BY?
- 17.** ¿Cuál palabra clave deberá agregarse a una cláusula ORDER BY para clasificar los datos en orden descendente?

Capítulo 8

Modificar datos SQL

Habilidades y conceptos clave

- Insertar datos SQL
 - Actualizar datos SQL
 - Eliminar datos SQL
-

Una de las funciones principales de cualquier base de datos es la capacidad de manejar los datos almacenados dentro de sus tablas. Los usuarios designados deben ser capaces de insertar, actualizar y eliminar los datos según sea necesario para mantener el flujo de la base de datos y asegurarse que sólo los datos adecuados están siendo almacenados. SQL proporciona tres instrucciones para el manejo básico de datos: INSERT, UPDATE y DELETE. En este capítulo se examinarán cada una de estas instrucciones y se demostrará cómo pueden ser utilizadas en un ambiente SQL para modificar los datos de la base de datos.

Insertar datos SQL

En el capítulo 7, el ejercicio 7-1 contenía una breve introducción acerca de la instrucción INSERT. Como se puede ver en ese ejercicio, la instrucción INSERT permite agregar datos a las diferentes tablas en una base de datos. Se presenta la sintaxis básica en esta sección y una sintaxis alternativa en la siguiente sección (“Insertar valores desde una instrucción SELECT”). La sintaxis de una instrucción INSERT básica es relativamente sencilla:

```
INSERT INTO <nombre de la tabla>
[ ( <nombre de la columna> ) [ { , <nombre de la columna> }... ] ]
VALUES ( <valor> [ { , <valor> }... ] )
```

En la sintaxis, sólo son obligatorias la primera y la última líneas. La segunda línea es opcional. Tanto la primera como la segunda línea son parte de la cláusula INSERT INTO. En esta cláusula se debe identificar el nombre de la tabla (o vista) en la cual se agregarán los datos. El nombre de la tabla sigue a las palabras clave INSERT INTO. Entonces el usuario tiene la opción de identificar los nombres de la columna en la tabla que recibirán los datos. Éste es el propósito de la segunda línea en la sintaxis. Es posible especificar una o más columnas, y todas ellas deberán estar dentro de paréntesis. Si se especifican más columnas, éstas deberán estar separadas por comas.

NOTA

La mayoría de las implementaciones de SQL soportan las vistas INSERT INTO. Sin embargo, existen restricciones. Por ejemplo, no es posible insertar INSERT INTO en una vista si existen columnas de la tabla que no están incluidas en la vista y que no permiten valores nulos ni tienen un valor por defecto definido. Adicionalmente, si la vista tiene más de una tabla base, podría no ser posible insertar INSERT INTO en lo absoluto, y si se pudiera, se le requeriría al usuario nombrar columnas desde sólo una de las tablas base debido a que un INSERT sólo puede afectar a una tabla base. Siempre revise la documentación del fabricante para mayor información.

En la tercera línea de la sintaxis, que se refiere a la cláusula **VALUES**, se deben especificar uno o más valores que serán ingresados en la tabla. La lista de valores deberá estar encerrada en paréntesis y, si se especifica más de uno, deberán estar separados utilizando comas. Asimismo, los valores deben cumplir los siguientes requisitos:

- Si los nombres de columna no se especifican en la cláusula **INSERT INTO**, entonces deberá haber un valor por cada columna en la tabla y los valores deberán estar en el mismo orden en el que están definidos en la tabla.
- Si los nombres de columna se especifican en la cláusula **INSERT INTO**, entonces deberá haber exactamente un valor por cada columna especificada y esos valores deberán estar en el mismo orden en el que están definidos en la cláusula **INSERT INTO**. Sin embargo, los nombres y valores de columna no tienen que estar en el mismo orden que las columnas en la definición de la tabla.
- Se debe proporcionar un valor por cada columna en la tabla *excepto* para las columnas que permiten valores nulos o que tienen un valor definido por defecto.
- Cada valor con un carácter del tipo de datos de cadena debe estar encerrado en comillas sencillas.
- Se puede utilizar la palabra clave **NULL** (o **null**) como el valor de los datos en la cláusula **VALUES** para asignar un valor nulo a cualquier columna que permita valores nulos.

NOTA

Muchos programadores de SQL prefieren especificar los nombres de columna dentro de la cláusula **INSERT INTO**, incluso si no es necesario, debido a que proporciona un método para documentar cuáles columnas suponen recibir los datos. Esta práctica también hace a la instrucción **INSERT** menos propensa a errores y a otros problemas provocados al añadir nuevas columnas o al cambiar el orden de las columnas en algún momento futuro. Por estas razones, muchas organizaciones requieren el uso de los nombres de columna en todas las instrucciones **INSERT**.

Ahora veamos algunos ejemplos de la instrucción **INSERT**. Para estos ejemplos se utilizará la tabla **INVENTARIO_CD**. La tabla está basada en la siguiente definición de tabla:

```
CREATE TABLE INVENTARIO_CD
( NOMBRE_CD          VARCHAR(60)                NOT NULL,
  TIPO_MUSICA        VARCHAR(15),
  EDITOR             VARCHAR(50) DEFAULT 'Independiente' NOT NULL,
  EN_EXISTENCIA      INT                        NOT NULL );
```

En el primer ejemplo se mostrarán los valores insertados en cada columna en la tabla **INVENTARIO_CD**:

```
INSERT INTO INVENTARIO_CD
VALUES ( 'Patsy Cline: 12 Greatest Hits', 'Country', 'MCA Records', 32 );
```

Observe que la cláusula **INSERT INTO** incluye solamente el nombre de la tabla **INVENTARIO_CD**, pero no especifica ninguna columna. En la cláusula **VALUES** se han especificado cuatro valores. Los valores están separados por comas, y los valores con caracteres del tipo de datos de

cadena están encerrados en comillas sencillas. Si hacemos referencia a la definición de tabla, se verá que los valores especificados en la cláusula VALUES se encuentran en el mismo orden que las definiciones de columna.

Cuando se ejecuta la instrucción INSERT mostrada en el ejemplo, los datos son añadidos a la tabla INVENTARIO_CD, como se muestra en la figura 8-1.

Si se intentara ejecutar una instrucción INSERT como en el último ejemplo, pero incluyendo solamente tres valores en lugar de cuatro, se recibiría como resultado un error. Por ejemplo, no sería posible ejecutar la siguiente instrucción:

```
INSERT INTO INVENTARIO_CD
VALUES ( 'Patsy Cline: 12 Greatest Hits', 'MCA Records', 32 );
```

En este ejemplo se han especificado solamente tres valores. En este caso, el valor faltante es para la columna TIPO_MUSICA. A pesar de que esta columna acepta valores nulos, la implementación de SQL no tiene forma de saber cuál valor está siendo omitido, y por lo tanto se arroja un error.

En lugar de dejar el valor fuera de la cláusula VALUES, se puede especificar un valor nulo, como se muestra en el siguiente ejemplo:

```
INSERT INTO INVENTARIO_CD
VALUES ( 'Out of Africa', null, 'MCA Records', 29 );
```

Si se ejecuta la instrucción INSERT, la tabla no incluirá una fila adicional. La figura 8-2 muestra cómo se vería la tabla, asumiendo que las dos instrucciones INSERT han sido ejecutadas.

El valor nulo se ingresó en la columna TIPO_MUSICA, y los otros valores fueron ingresados en sus columnas apropiadas. Si no se permitiera un valor nulo en la columna TIPO_MUSICA, se tendría que haber especificado un valor.

NOTA

La figura 8-2 muestra la nueva fila que está siendo insertada después de la fila existente en la tabla. Sin embargo, la fila podría ser insertada en cualquier lugar en una tabla, dependiendo de cómo inserte filas la implementación de SQL. El estándar de SQL no especifica dónde se insertará una fila en una tabla. De hecho, jamás se deberá confiar en que las filas en una tabla estén en un orden en particular (se deberá utilizar la cláusula ORDER BY cuando los resultados de SELECT necesiten estar en una secuencia en particular).

En lugar de proporcionar un valor para cada columna cuando se inserta una fila, se puede especificar cuáles columnas recibirán valores. Por ejemplo, se puede especificar los valores para las

NOMBRE_CD: VARCHAR(60)	TIPO_MUSICA: VARCHAR(15)	EDITOR: VARCHAR(50)	EN_EXISTENCIA: INT
Patsy Cline: 12 Greatest Hits	Country	MCA Records	32

Figura 8-1 La tabla INVENTARIO_CD con la nueva fila de datos.

NOMBRE_CD: VARCHAR(60)	TIPO_MUSICA: VARCHAR(15)	EDITOR: VARCHAR(50)	EN_EXISTENCIA INT
Patsy Cline: 12 Greatest Hits	Country	MCA Records	32
Out of Africa	NULL	MCA Records	29

Figura 8-2 La tabla INVENTARIO_CD con dos filas de datos.

columnas NOMBRE_CD, EDITOR y EN_EXISTENCIA de la tabla INVENTARIO_CD, como se muestra en el siguiente ejemplo:

```
INSERT INTO INVENTARIO_CD ( NOMBRE_CD, EDITOR, EN_EXISTENCIA )
VALUES ( 'Fundamental', 'Capitol Records', 34 );
```

En este caso, un solo valor ha sido especificado para cada una de las columnas identificadas dentro de la cláusula INSERT INTO, y los valores se especifican en el mismo orden que las columnas en la cláusula INSERT INTO. Observe que la instrucción INSERT no incluye a la columna TIPO_MUSICA dentro de la cláusula INSERT INTO ni de la cláusula VALUES. Se puede omitir esta columna debido a que los valores nulos se permiten en esa columna. Si se ejecutara esta instrucción, la tabla INVENTARIO_CD ahora tendría una tercera fila (mostrada la figura 8-3).

Una vez más, el valor nulo (NULL) es agregado automáticamente a la columna TIPO_MUSICA. Si un valor por defecto hubiera sido definido para la columna, ese valor habría sido agregado. Por ejemplo, la siguiente instrucción INSERT omite la columna EDITOR en lugar de la columna TIPO_MUSICA:

```
INSERT INTO INVENTARIO_CD ( NOMBRE_CD, TIPO_MUSICA, EN_EXISTENCIA )
VALUES ( 'Orlando', 'Soundtrack', 5 );
```

Cuando la fila es agregada a la tabla INVENTARIO_CD, el valor por defecto (Independiente) es agregado a la columna EDITOR, como se muestra la figura 8-4.

NOMBRE_CD: VARCHAR(60)	TIPO_MUSICA: VARCHAR(15)	EDITOR: VARCHAR(50)	EN_EXISTENCIA: INT
Patsy Cline: 12 Greatest Hits	Country	MCA Records	32
Out of Africa	NULL	MCA Records	29
Fundamental	NULL	Capitol Records	34

Figura 8-3 La tabla INVENTARIO_CD con tres filas de datos.

NOMBRE_CD: VARCHAR(60)	TIPO_MUSICA: VARCHAR(15)	EDITOR: VARCHAR(50)	EN_EXISTENCIA: INT
Patsy Cline: 12 Greatest Hits	Country	MCA Records	32
Out of Africa	NULL	MCA Records	29
Fundamental	NULL	Capitol Records	34
Orlando	Soundtrack	Independent	5

Figura 8-4 La tabla INVENTARIO_CD con cuatro filas de datos.

Si se intenta ejecutar una instrucción INSERT omitiendo una columna que no permita valores nulos y que no tenga un valor por defecto definido, se arrojará como resultado un error. Cuando se inserta una nueva fila, el RDBMS debe obtener un valor para cada columna desde algún lado, por lo que si no se permiten los valores nulos, entonces el valor debe venir ya sea de un valor por defecto (si está definido) o de la cláusula VALUES de la instrucción INSERT.

NOTA

Los valores que se especifiquen en la cláusula VALUES deberán hacerse de acuerdo con todas las restricciones localizadas en una tabla. Esto significa que los valores deben cumplir con los tipos de datos o dominios asociados con una columna. Además, los valores están confinados por cualquier restricción definida en la tabla. Por ejemplo, una restricción de clave foránea pudiera evitar que se agregue cualquier valor que viole la limitación, o una limitación de revisión pudiera limitar el rango de valores que pueden ser ingresados a la tabla. Asegúrese de conocer ampliamente las restricciones localizadas en una tabla antes de intentar agregar datos en ella. Se puede aprender más acerca de los tipos de datos en el capítulo 3. Se puede aprender más acerca de dominios y restricciones en el capítulo 4.

Desde luego, también es posible especificar todas las columnas en la cláusula INSERT INTO. Si se realiza esto, debe especificarse correctamente el mismo número de valores, en el mismo orden en el cual están especificadas las columnas. La siguiente instrucción INSERT ingresa valores a todas las columnas de la tabla INVENTARIO_CD:

```
INSERT INTO INVENTARIO_CD ( NOMBRE_CD, MUSIC_TYPE, EDITOR, EN_EXISTENCIA )
VALUES ( 'Court and Spark', 'Pop', 'Asylum Records', 22 );
```

Cuando se ejecuta esta instrucción, una fila es agregada a la tabla INVENTARIO_CD, con un valor para cada columna. La figura 8-5 muestra la nueva fila, junto con las cuatro filas anteriores que se habían insertado. Si se omitiera uno de los valores desde la cláusula VALUES (incluso cuando en esa columna relacional se permitan los valores nulos) se recibiría un mensaje de error cuando se ejecute la instrucción.

Insertar valores desde una instrucción SELECT

Anteriormente en este capítulo, al inicio de la sección “Insertar datos SQL”, se mencionó que la cláusula VALUES es obligatoria y que es necesario especificar por lo menos un valor. Sin embar-

NOMBRE_CD: VARCHAR(60)	TIPO_MUSICA: VARCHAR(15)	EDITOR: VARCHAR(50)	EN_EXISTENCIA: INT
Patsy Cline: 12 Greatest Hits	Country	MCA Records	32
Out of Africa	NULL	MCA Records	29
Fundamental	NULL	Capitol Records	34
Orlando	Soundtrack	Independent	5
Court and Spark	Pop	Asylum Records	22

Figura 8-5 La tabla INVENTARIO_CD con cinco filas de datos.

go, existe una alternativa a la cláusula VALUES. Se puede utilizar una instrucción SELECT para especificar los valores que se quieran ingresar en una tabla. La clave para utilizar una instrucción SELECT, al igual que al utilizar una cláusula VALUES, es asegurarse que el número de valores aplicados por la instrucción SELECT coincida con el número requerido de valores, y que éstos cumplan con cualquier restricción de la tabla correspondiente. Observemos un ejemplo.

Supongamos que, además de la tabla INVENTARIO_CD que se ha estado utilizando en los ejemplos anteriores, la base de datos también tuviera una segunda tabla llamada INVENTARIO_CD_2, que incluyera dos columnas, como se muestra en la siguiente definición de tabla:

```
CREATE TABLE INVENTARIO_CD_2
( NOMBRE_CD_2      VARCHAR(60)  NOT NULL,
  EN_EXISTENCIA_2  INT           NOT NULL );
```

La columna NOMBRE_CD_2 en la tabla INVENTARIO_CD_2 tiene los mismos tipos de datos que la columna NOMBRE_CD en la tabla INVENTARIO_CD, y la columna EN_EXISTENCIA_2 en la tabla INVENTARIO_CD_2 tiene los mismos tipos de datos que la columna EN_EXISTENCIA en la tabla INVENTARIO_CD. Como resultado, los valores tomados de las dos columnas en una tabla pueden ser insertados en las dos columnas de la segunda tabla.

NOTA

Una columna en una tabla no tiene que contener los mismos tipos de datos que una columna en otra tabla para que los valores sean copiados de una a la otra; basta con que los valores insertados en la tabla destino cumplan con las restricciones de datos de esa tabla.

Al utilizar la instrucción INSERT, se pueden copiar valores de la tabla INVENTARIO_CD a la tabla INVENTARIO_CD_2. La siguiente instrucción INSERT incluye una instrucción SELECT que consulta a la tabla INVENTARIO_CD:

```
INSERT INTO INVENTARIO_CD_2
SELECT NOMBRE_CD, EN_EXISTENCIA
FROM INVENTARIO_CD;
```

NOMBRE_CD: VARCHAR(60)	EN_EXISTENCIA: INT
Patsy Cline: 12 Greatest Hits	32
Out of Africa	29
Fundamental	34
Orlando	5
Court and Spark	22

Figura 8-6 La tabla INVENTARIO_CD_2 con cinco filas de datos.

Como se puede ver, ninguna columna está especificada en la cláusula INSERT INTO; como resultado, los valores serán insertados en ambas columnas en la tabla INVENTARIO_CD_2. En la segunda línea de la instrucción se utiliza una instrucción SELECT para tomar los valores desde las columnas NOMBRE_CD y EN_EXISTENCIA de la tabla INVENTARIO_CD. Los valores serán entonces insertados en sus columnas respectivas en la tabla INVENTARIO_CD_2, como se muestra la figura 8-6.

Observe que la tabla INVENTARIO_CD_2 contiene las mismas cinco filas de datos que se muestran en la figura 8-5, sólo que la tabla INVENTARIO_CD_2 contiene únicamente dos columnas: NOMBRE_CD_2 e EN_EXISTENCIA_2.

Al igual que cualquier otra instrucción SELECT, la instrucción SELECT que se utiliza en una instrucción INSERT puede contener una cláusula WHERE. En la siguiente instrucción INSERT, la instrucción SELECT contiene una cláusula WHERE que limita los valores de EN_EXISTENCIA a una cantidad mayor a 10:

```
INSERT INTO INVENTARIO_CD_2
  SELECT NOMBRE_CD, EN_EXISTENCIA
  FROM INVENTARIO_CD
  WHERE EN_EXISTENCIA > 10;
```

Si se ejecutara esta instrucción, solamente cuatro filas serían añadidas a la tabla INVENTARIO_CD_2, en lugar de las cinco filas que se vieron en el ejemplo anterior. En este caso, la cláusula WHERE funciona exactamente igual que una cláusula WHERE en cualquier instrucción SELECT. Como resultado, cualquier fila con un valor EN_EXISTENCIA que no sea mayor a 10 es eliminado de los resultados de la consulta. Esos nuevos resultados filtrados se insertan entonces en la tabla INVENTARIO_CD_2.

Actualizar datos SQL

Como su nombre lo indica, la instrucción UPDATE permite actualizar los datos en una base de datos SQL. Con la instrucción UPDATE se pueden modificar datos en una o más filas para una o más columnas. La sintaxis para la instrucción UPDATE se puede mostrar de la manera siguiente:

```
UPDATE <nombre de la tabla>
SET <determinar expresión de la cláusula> [ {, <determinar expresión de la cláusula> }... ]
[ WHERE <condición de búsqueda> ]
```

Como se puede ver, la cláusula UPDATE y la cláusula SET son obligatorias, mientras que la cláusula WHERE es opcional. En la cláusula UPDATE se debe especificar el nombre de la tabla (o vista) que se está actualizando. En la cláusula SET se debe especificar una o más expresiones de cláusula, lo cual se discutirá con más detalle posteriormente en este capítulo. En la cláusula WHERE, al igual que con la cláusula WHERE en una instrucción SELECT (véase el capítulo 7), se debe especificar una condición de búsqueda. La cláusula WHERE funciona aquí de una forma muy parecida a como lo hace en la instrucción SELECT. Se especifica una condición o conjunto de condiciones que actúa como un filtro para las filas que se están actualizando. Solamente las filas que cumplen con estas condiciones son actualizadas. En otras palabras, solamente las filas cuyo resultado es verdadero son actualizadas.

NOTA

SQL permite utilizar nombres de vistas en las instrucciones UPDATE. Sin embargo, si la vista está basada en múltiples tablas, todas las columnas que se están actualizando deben venir desde una sola tabla base, y no debe haber otras restricciones como se describe en su documentación de DBMS.

Ahora regresemos a la cláusula SET. Como se puede ver, la cláusula incluye el marcador de posición < determinar expresión de la cláusula >. Se deben especificar una o más determinar expresión de la cláusula. Si se especifica más de una, se deben separar utilizando comas. La sintaxis del marcador de posición < determinar expresión de la cláusula > puede descomponerse como se muestra a continuación:

< nombre de la columna > = < expresión de valor >

Básicamente, se debe especificar un nombre de columna (desde la tabla que se está actualizando) y proporcionar un valor que el valor en la columna deberá igualar. Por ejemplo, supongamos que se requiere que un valor en la columna EN_EXISTENCIA sea cambiado a 37. (Sin importar cuál sea el valor actual en esa columna.) Determinar expresión de la cláusula quedaría como sigue: EN_EXISTENCIA = 37. En este caso, la expresión de valor es 37; sin embargo, la expresión de valor puede ser más complicada que eso. Por ejemplo, se puede basar el nuevo valor en un valor antiguo: EN_EXISTENCIA = (EN_EXISTENCIA + 1). En este caso, la expresión de valor es EN_EXISTENCIA + 1, la cual agrega 1 al valor actual en la columna EN_EXISTENCIA para darle un nuevo valor. En este ejemplo, si el valor original era 37, el nuevo valor será 38.

Ahora que se ha dado un vistazo a las diferentes partes de la instrucción UPDATE, pongámonos todas juntas utilizando algunos ejemplos. Los ejemplos que se utilizarán estarán basados en la tabla INVENTARIO_CD, que se muestra en la figura 8-5.

En el primer ejemplo se utiliza la instrucción UPDATE para cambiar los valores de la columna EN_EXISTENCIA a 27, como se muestra en la siguiente instrucción SQL:

```
UPDATE INVENTARIO_CD
SET EN_EXISTENCIA = 27;
```

Esta instrucción realiza exactamente lo que se pudiera esperar: cambia cada fila en la tabla INVENTARIO_CD para que la columna EN_EXISTENCIA contenga un valor de 27 para cada fila. El resultado es óptimo si eso es lo que se quería, pero es poco probable que se quiera cambiar cada fila en una tabla para que todos los valores de esa columna sean iguales en cada fila. Es mucho más probable que se quiera precisar la actualización utilizando una cláusula WHERE.

En el siguiente ejemplo se modifica la instrucción UPDATE anterior para incluir una cláusula WHERE:

```
UPDATE INVENTARIO_CD
  SET EN_EXISTENCIA = 27
  WHERE NOMBRE_CD = 'Out of Africa';
```

La instrucción UPDATE aún cambia la columna EN_EXISTENCIA a un valor de 27, pero sólo lo hace para las filas que concuerdan con la condición de búsqueda en la cláusula WHERE. En este caso, solamente una fila cumple con esa condición: Out of Africa.

Algunas veces también se requiere cambiar un valor basado en un valor que ya exista, por ejemplo, la cantidad de inventario en existencia. Por ejemplo, se pueda añadir 2 al valor en la columna EN_EXISTENCIA:

```
UPDATE INVENTARIO_CD
  SET EN_EXISTENCIA = (EN_EXISTENCIA + 2)
  WHERE NOMBRE_CD = 'Out of Africa';
```

Si la fila Out of Africa contiene el valor 27 en la columna EN_EXISTENCIA, y se ejecuta esta instrucción UPDATE, el nuevo valor será 29. Si se ejecuta esta instrucción sin la cláusula WHERE, el valor 2 será agregado al valor EN_EXISTENCIA para cada fila en la tabla.

La cláusula WHERE también permite especificar más de un predicado, al igual que se puede hacer con una cláusula WHERE en la instrucción SELECT. En el siguiente ejemplo, se sustrae 2 del valor EN_EXISTENCIA para cualquier fila que contenga un valor TIPO_MUSICA de Country y un valor EN_EXISTENCIA mayor a 30:

```
UPDATE INVENTARIO_CD
  SET EN_EXISTENCIA = (EN_EXISTENCIA - 2)
  WHERE TIPO_MUSICA = 'Country' AND EN_EXISTENCIA > 30;
```

Solamente una fila (Patsy Cline: 12 Greatest Hits) cumple con las condiciones de búsqueda especificadas en la cláusula WHERE. El valor EN_EXISTENCIA para esa fila ha sido cambiado de 32 a 30.

También es posible especificar múltiples expresiones en la cláusula SET. En otras palabras, se pueden cambiar los valores de más de una columna al mismo tiempo. Por ejemplo, supongamos que se quiere cambiar el valor EDITOR y el valor EN_EXISTENCIA de la fila Orlando. La instrucción UPDATE necesaria podría lucir de la siguiente forma:

```
UPDATE INVENTARIO_CD
  SET PUBLISHER = 'Sarabande Records',
      EN_EXISTENCIA = (EN_EXISTENCIA * 2)
  WHERE NOMBRE_CD = 'Orlando';
```

Observe que las dos expresiones en la cláusula SET están separadas por una coma. Cuando se ejecuta esta instrucción, el valor EDITOR se cambia de Independiente a Sarabande Records, y el valor EN_EXISTENCIA se cambia de 5 a 10. (El valor 5 se multiplica por 2.)

Algo que no se puede realizar, sin embargo, es cambiar el valor para la misma columna para dos diferentes filas si se está tratando de dar diferentes valores para esas filas. Veamos un ejemplo para hacer esto más claro. Supongamos que se quiere actualizar el valor para la fila Out of Africa y para la fila Fundamental, pero se quiere actualizar estas filas con diferentes valores. La fila Out of Africa debe contener el valor Soundtrack para TIPO_MUSICA, y la fila Fundamental debe con-

tener el valor Blues para TIPO_MUSICA. Como resultado, se puede intentar ejecutar una instrucción similar a la siguiente:

```
UPDATE INVENTARIO_CD
    SET TIPO_MUSICA = 'Soundtrack',
        TIPO_MUSICA = 'Blues'
WHERE NOMBRE_CD = 'Out of Africa' OR NOMBRE_CD = 'Fundamental';
```

Si se intenta ejecutar esta instrucción, la implementación de SQL no sabrá cuál valor de TIPO_MUSICA debe poner en qué fila, y arrojará un error. Para manejar una situación como ésta, se necesitarán crear dos instrucciones UPDATE separadas:

```
UPDATE INVENTARIO_CD
    SET TIPO_MUSICA = 'Soundtrack'
WHERE NOMBRE_CD = 'Orlando';
UPDATE INVENTARIO_CD
    SET TIPO_MUSICA = 'Blues'
WHERE NOMBRE_CD = 'Fundamental';
```

Actualizar valores desde una instrucción SELECT

En la sección “Insertar valores desde una instrucción SELECT” ya presentada en este capítulo, se dijo que se puede utilizar una instrucción SELECT en lugar de una cláusula VALUES. También se puede utilizar una instrucción SELECT en la cláusula SET de la instrucción UPDATE. La instrucción SELECT arroja el valor que está definido en la porción <expresión de valor> de la expresión determinar cláusula. En otras palabras, la instrucción SELECT se agrega a la derecha del signo de igual.

Pongamos algunos ejemplos para ver cómo funciona esto. Los siguientes ejemplos están basados en los datos originales de la tabla INVENTARIO_CD (mostrada en la figura 8-5) y la tabla INVENTARIO_CD_2 (mostrada en la figura 8-6). Supongamos que se quieren actualizar los datos en la tabla INVENTARIO_CD_2 utilizando valores de la tabla INVENTARIO_CD. Se puede crear una instrucción UPDATE que sea similar a la siguiente:

```
UPDATE INVENTARIO_CD_2
    SET EN_EXISTENCIA_2 =
        ( SELECT AVG(EN_EXISTENCIA)
          FROM INVENTARIO_CD );
```

La instrucción SELECT calcula el promedio de los valores EN_EXISTENCIA en la tabla INVENTARIO_CD, el cual es 24, por lo que determinar la expresión de la cláusula puede ser interpretado como sigue: EN_EXISTENCIA_2 = 24. Como resultado, todos los valores EN_EXISTENCIA_2 en la tabla INVENTARIO_CD_2 son determinados a 24. Desde luego, probablemente no se requiere que todos los valores de EN_EXISTENCIA_2 sean iguales, por lo que puede limitarse cuáles filas serán actualizadas al agregar una cláusula WHERE a la instrucción UPDATE:

```
UPDATE INVENTARIO_CD_2
    SET EN_EXISTENCIA_2 =
        ( SELECT AVG(EN_EXISTENCIA)
          FROM INVENTARIO_CD )
WHERE NOMBRE_CD_2 = 'Orlando';
```

Ahora, solamente la fila Orlando será actualizada y el valor de EN_EXISTENCIA_2 será cambiado a 24.

Incluso se puede agregar una cláusula WHERE a la instrucción SELECT, como se muestra en el siguiente ejemplo:

```
UPDATE INVENTARIO_CD_2
SET EN_EXISTENCIA_2 =
    ( SELECT EN_EXISTENCIA
      FROM INVENTARIO_CD
      WHERE NOMBRE_CD = 'Orlando' )
WHERE NOMBRE_CD_2 = 'Orlando';
```

En este caso, el valor EN_EXISTENCIA de 5 se toma directamente desde la fila Orlando de la tabla INVENTARIO_CD y se utiliza como la porción <expresión de valor> de la expresión determinar cláusula. Como resultado, determinar la expresión de la cláusula puede ser interpretado como lo siguiente: EN_EXISTENCIA_2 = 5. (Por supuesto, el valor en la tabla INVENTARIO_CD_2 no cambiará debido a que ya es 5, pero si tuviera cualquier otro valor diferente, éste se habría actualizado a 5.)

Se puede agregar una capa más de complejidad a la instrucción UPDATE modificando la cláusula SET todavía más. Por ejemplo, supongamos que se quiere incrementar el valor en 2 antes de insertarlo en la columna EN_EXISTENCIA_2. Para hacer eso, se puede cambiar la expresión de valor a la siguiente:

```
UPDATE INVENTARIO_CD_2
SET EN_EXISTENCIA_2 =
    ( SELECT EN_EXISTENCIA
      FROM INVENTARIO_CD
      WHERE NOMBRE_CD = 'Orlando' ) + 2
WHERE NOMBRE_CD_2 = 'Orlando';
```

Una vez más, la cláusula SELECT toma el valor de 5 desde la columna EN_EXISTENCIA de la tabla INVENTARIO_CD, pero esta vez el valor 2 es agregado al valor arrojado por la instrucción SELECT, resultando un total de 7. Como resultado, el nuevo determinar la expresión de la cláusula puede ser representado así: EN_EXISTENCIA_2 = (5) + 2. Si se ejecutara esta instrucción, el valor de EN_EXISTENCIA_2 cambiaría a 7 en la fila Orlando de la tabla INVENTARIO_CD_2.

Al combinar la cláusula SET con la cláusula WHERE, se pueden crear instrucciones UPDATE que puedan calcular valores muy específicos que se utilicen para modificar cualquier número de filas y columnas que se necesite actualizar. Sin embargo, al igual que con la instrucción INSERT, cualquier valor que se modifique deberá cumplir con las restricciones de la tabla. En otras palabras, los nuevos valores deben regirse por los tipos de datos, dominios y limitaciones aplicables.

Eliminar datos SQL

De todas las instrucciones para modificación de datos soportadas por SQL, probablemente la instrucción DELETE sea la más sencilla. Contiene solamente dos cláusulas, y una sola de ellas es obligatoria. La siguiente sintaxis muestra qué tan básica es la instrucción DELETE:

```
DELETE FROM <nombre de la tabla>
[ WHERE <condición de búsqueda> ]
```

Como se puede ver, la cláusula DELETE FROM requiere que se especifique el nombre de la tabla (o vista) de la cual se quieren eliminar filas. La cláusula WHERE, la cual es similar a la cláusula WHERE en una instrucción SELECT y en una instrucción UPDATE, requiere que se especifique una condición de búsqueda. Si no se incluye una cláusula WHERE en la instrucción DELETE, todas las filas serán eliminadas de la tabla especificada. Es importante comprender que la instrucción DELETE no elimina la tabla en sí, sino solamente filas en la tabla (la instrucción DROP TABLE, como se describió en el capítulo 3, se utiliza para eliminar definiciones de tabla de la base de datos).

NOTA

SQL soporta vistas de referencia en la instrucción DELETE, pero la eliminación real se realiza en filas en la tabla base. Prácticamente en ninguna de las implementaciones se pueden eliminar filas utilizando vistas que hagan referencia a más de una tabla base (véase la documentación del fabricante para detalles más específicos).

Observe que en la instrucción DELETE no hay nombres de columna especificados. Esto se debe a que no es posible eliminar valores de columna individuales desde una tabla. Solamente se pueden eliminar filas. Si se necesita eliminar un valor de columna específico, se deberá utilizar una instrucción UPDATE para determinar el valor a nulo. Pero sólo es posible realizar esto si en esa columna se soportan valores nulos.

Ahora veamos un par de ejemplos de la instrucción DELETE. El primer ejemplo elimina todos los datos (todas las filas) de la tabla INVENTARIO_CD, mostrado en la figura 8-5:

```
DELETE FROM INVENTARIO_CD;
```

Eso es todo lo que se necesita. Desde luego, deberá utilizarse esta instrucción solamente si se quieren eliminar *todos* los datos de INVENTARIO_CD. Aunque en ocasiones se puede ejecutar esta instrucción donde sea necesario eliminar cada fila de una tabla, es un poco más probable que se prefiera utilizar la cláusula WHERE, para especificar cuáles filas serán eliminadas. Se puede modificar la instrucción que acabamos de ver para eliminar solamente filas donde el valor TIPO_MUSICA sea Country:

```
DELETE FROM INVENTARIO_CD
WHERE TIPO_MUSICA = 'Country';
```

Cuando se ejecuta esta instrucción, todas las filas cuyo valor TIPO_MUSICA es Country serán eliminadas de la tabla INVENTARIO_CD, que en este caso es la fila Patsy Cline: 12 Greatest Hits.

Ahora modificamos la instrucción DELETE un poco más incluyendo dos predicados en la cláusula WHERE:

```
DELETE FROM INVENTARIO_CD
WHERE TIPO_MUSICA = 'Pop'
OR EDITOR = 'Independiente';
```

Esta instrucción eliminará cualquier fila en la tabla CD_INVENTORY que incluya a un valor TIPO_MUSICA de Pop o un valor EDITOR de Independiente, lo que significa que la fila Court and Spark y la fila Orlando serán eliminadas.

Como se puede ver, el número de filas que se eliminan de una tabla depende de las condiciones de búsqueda definidas dentro de la cláusula WHERE. Cuando no se especifica una cláusula

WHERE, todas las filas que arrojan un valor verdadero son eliminadas de la tabla. La cláusula WHERE permite especificar exactamente cuáles filas deberán ser eliminadas de la tabla.

Pruebe esto 8-1 Modificar datos SQL

En este ejercicio se utilizarán todas las instrucciones de modificación de datos analizadas en este capítulo para cambiar los datos en la base de datos INVENTARIO. Se utilizará la instrucción INSERT para agregar datos, la instrucción UPDATE para modificar los datos, y la instrucción DELETE para eliminar los datos de la base de datos. Debido a que se estará trabajando solamente con datos, no se afectará la estructura subyacente de las tablas. Se puede descargar el archivo Try_This_08.txt (en inglés), que contiene las instrucciones SQL utilizadas en este ejercicio.

Paso a paso

1. Abra la aplicación cliente para su RDBMS y conéctese con la base de datos INVENTARIO.
2. Primero se añadirá una nueva compañía a la tabla DISQUERAS_CD. La compañía es DRG Records y tendrá un valor ID_DISQUERA de 837. Ingrese y ejecute la siguiente instrucción SQL:

```
INSERT INTO DISQUERAS_CD
VALUES ( 837, 'DRG Records' );
```

Una fila se añadirá a la tabla DISQUERAS_CD.

3. Ahora agreguemos un nuevo CD a la tabla DISCOS_COMPACTOS. El CD tiene el nombre *Ann Hampton Callaway*, el cual tiene un valor ID_DISCO_COMPACTO de 116. Se encuentran 14 de estos CD en existencia y el valor ID_DISQUERA deberá ser 836. (Éste no es el valor ID_DISQUERA correcto, pero se utilizará aquí para propósitos de este ejercicio.) Ingrese y ejecute la siguiente instrucción SQL:

```
INSERT INTO DISCOS_COMPACTOS
VALUES ( 116, 'Ann Hampton Callaway', 836, 14 );
```

Una fila se añadirá a la tabla DISCOS_COMPACTOS. El valor ID_DISQUERA de 836 representa Sarabande Records.

4. Ahora agregaremos otra fila en la tabla DISCOS_COMPACTOS; sólo que esta vez, la instrucción INSERT especificará los nombres de las columnas de la tabla destino. Se insertará un CD llamado *Rhythm Country and Blues*. La nueva fila contendrá un valor ID_DISCO_COMPACTO de 117, un valor ID_DISQUERA de 832 (MCA Records) y un valor EN_EXISTENCIA de 21. Ingrese y ejecute la siguiente instrucción SQL:

```
INSERT INTO DISCOS_COMPACTOS
( ID_DISCO_COMPACTO, TITULO_CD, ID_DISQUERA, EN_EXISTENCIA )
VALUES ( 117, 'Rhythm Country and Blues', 832, 21 );
```

Una fila se añadirá a la tabla DISCOS_COMPACTOS.

5. Después de ingresar la fila Rhythm Country and Blues, se puede observar que el valor EN_EXISTENCIA es incorrecto y que es necesario actualizar ese valor a 25. Ingrese y ejecute la siguiente instrucción SQL:

```
UPDATE DISCOS_COMPACTOS
    SET EN_EXISTENCIA = 25
    WHERE ID_DISCO_COMPACTO = 117;
```

El valor EN_EXISTENCIA de la fila Rhythm Country and Blues será cambiado a 25.

6. Ahora se puede observar que se ingresó el valor incorrecto ID_DISQUERA para la fila Ann Hampton Callaway. Sin embargo, se quiere modificar el valor existente especificando el nombre de la compañía en lugar del valor ID_DISQUERA. El nombre de la compañía es DRG Records, que se añadió a la tabla DISQUERAS_CD en el paso 2. Ingrese y ejecute la siguiente instrucción SQL:

```
UPDATE DISCOS_COMPACTOS
    SET ID_DISQUERA =
        ( SELECT ID_DISQUERA
          FROM DISQUERAS_CD
          WHERE NOMBRE_COMPAÑIA = 'DRG Records' )
    WHERE ID_DISCO_COMPACTO = 116;
```

En esta instrucción se utilizó la instrucción SELECT para tomar el valor ID_DISQUERA de la tabla DISQUERAS_CD. La instrucción arrojó un valor de 837. El valor 837 fue entonces utilizado como el valor ID_DISQUERA para la tabla DISCOS_COMPACTOS. Observe que no hubiera sido posible ingresar el valor de 837 en la columna ID_DISQUERA de la tabla DISCOS_COMPACTOS si no hubiera existido antes en la tabla DISQUERAS_CD. No sólo sucede esto debido a que una instrucción SELECT fue utilizada para tomar el valor, sino también debido a que la columna ID_DISQUERA en la tabla DISCOS_COMPACTOS es una clave externa que hace referencia a la tabla DISQUERAS_CD. Como resultado, el valor debe existir en la tabla referenciada antes de que pueda ser agregado a la tabla que hace referencia a ella. Véase el capítulo 4 para mayor información acerca de estas claves externas.

7. Ahora echemos un vistazo a los datos que se han ingresado y actualizado. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT *
    FROM DISCOS_COMPACTOS
    WHERE ID_DISCO_COMPACTO = 116
       OR ID_DISCO_COMPACTO = 117;
```

La instrucción SELECT pide los datos de todas las columnas en la tabla DISCOS_COMPACTOS, pero sólo para aquellas filas que tengan un valor ID_DISCO_COMPACTO de 116 o 117. Los resultados de la consulta deberán incluir dos filas. Verifique que la información en esas filas esté correcta. La fila Ann Hampton Callaway deberá tener un valor ID_DISQUERA de 837 y un valor EN_EXISTENCIA de 14, y la fila Rhythm Country and Blues deberá tener un valor ID_DISQUERA de 832 y un valor EN_EXISTENCIA de 25.

(continúa)

8. Ahora elimine las dos filas que se agregaron a la tabla DISCOS_COMPACTOS. Ingrese y ejecute la siguiente instrucción SQL:

```
DELETE FROM DISCOS_COMPACTOS
WHERE ID_DISCO_COMPACTO = 116
      OR ID_DISCO_COMPACTO = 117;
```

La fila Ann Hampton Callaway y Rhythm Country and Blues deberán haber sido eliminadas de la tabla DISCOS_COMPACTOS.

9. Por último elimine la fila que se añadió a la tabla DISQUERAS_CD. Ingrese y ejecute la siguiente instrucción SQL:

```
DELETE FROM DISQUERAS_CD
WHERE ID_DISQUERA = 837;
```

La fila DRG Records deberá haber sido eliminada de la tabla DISQUERAS_CD.

NOTA

Si se hubiera intentado eliminar esta fila anteriormente eliminando la fila Ann Hampton Callaway en la tabla DISCOS_COMPACTOS, se habría recibido un error debido a que el valor ID_DISQUERA en DISCOS_COMPACTOS hace referencia a la fila DRG Records en DISQUERAS_CD. La fila Ann Hampton Callaway tenía que ser eliminada primero, o el valor ID_DISQUERA tenía que ser cambiado a otro valor que cumpliera con la limitante de la clave externa.

10. Cierre la aplicación cliente.

Resumen de Pruebe esto

En este ejercicio se agregó una fila a la tabla ID_DISQUERA y dos filas a la tabla DISCOS_COMPACTOS. Luego se actualizaron las dos filas en la tabla DISCOS_COMPACTOS. Después de eso se eliminaron todas las filas que se habían creado. Al momento de terminar el ejercicio, la base de datos INVENTARIO deberá contener lo mismo que cuando se empezó el ejercicio. Como se puede ver, modificar los datos dentro de las tablas es un proceso bastante simple; sin embargo, las instrucciones de modificación de datos individuales pueden volverse mucho más complejas. Cuando aprenda técnicas más avanzadas para consultas de datos, usted será capaz de refinar sus instrucciones a un grado superior para que le proporcionen mayor flexibilidad al insertar, actualizar y eliminar datos.

Autoexamen Capítulo 8

1. ¿Cuál instrucción SQL deberá utilizarse para agregar datos a una tabla?
- A** SELECT
 - B** INSERT
 - C** UPDATE
 - D** DELETE

2. ¿Cuáles dos cláusulas son obligatorias en una instrucción INSERT?
3. ¿En cuál cláusula dentro de la instrucción INSERT se identifica la tabla que recibirá los nuevos datos?
4. Se crea la siguiente instrucción INSERT para agregar datos a la tabla ARTISTAS_INTERPRETES:

```
INSERT INTO ARTISTAS_INTERPRETES VALUES ( 12, 'Frank Sinatra' );
```

La tabla ARTISTAS_INTERPRETES incluye tres columnas. ¿Qué sucederá cuando se intente ejecutar esta instrucción?

5. ¿Qué información se debe especificar en la cláusula VALUES de una instrucción INSERT?
6. ¿Qué requerimientos deberán ser cumplidos por los valores en una cláusula VALUES?
7. Se está creando una instrucción INSERT para introducir datos en la tabla TIPOS_ARTISTA. La tabla incluye solamente dos columnas: ID_ART y NOMBRE_TIPO. Se quiere insertar una fila que incluye el valor ID_ART de 27 y el valor NOMBRE_TIPO de Gospel. ¿Cuál instrucción SQL deberá ser utilizada?
8. Se está creando una instrucción INSERT que inserta valores tomados desde otra tabla. ¿Qué tipo de instrucción o cláusula se puede utilizar en lugar de la cláusula VALUES para tomar datos desde la otra tabla?
 - A SELECT
 - B SET
 - C SELECT
 - D WHERE
9. ¿Cuál instrucción deberá utilizarse para modificar los datos existentes en una o más filas en una tabla?
 - A SELECT
 - B INSERT
 - C UPDATE
 - D DELETE
10. ¿Cuáles cláusulas son obligatorias en una instrucción UPDATE?
11. ¿Cuál es el propósito de la cláusula WHERE en una instrucción UPDATE?
12. Se está creando una instrucción UPDATE para actualizar los datos en la tabla ARTISTAS_INTERPRETES. Se quiere actualizar el valor ID_ART en la fila que contenga el valor ID_ART_INTER de 139. El nuevo valor ID_ART es 27. ¿Cuál instrucción SQL deberá ser utilizada?
13. Se está creando una instrucción UPDATE para actualizar los datos en la tabla ARTISTAS_INTERPRETES. Se quiere actualizar el valor ID_ART de cada fila a 27. ¿Cuál instrucción SQL deberá ser utilizada?

- 14.** Se están actualizando dos columnas en la tabla INVENTARIO_CD. Se quiere cambiar el valor EDITOR a MCA Records y se quiere duplicar el valor EN_EXISTENCIA. ¿Cuál cláusula SET deberá ser utilizada?
- 15.** Se está creando una instrucción UPDATE que incluye una cláusula SET con una expresión de valor. Se requiere que la expresión de valor tome un valor desde otra tabla en la base de datos. ¿Cuál instrucción o cláusula se puede utilizar como una expresión de valor para seleccionar datos desde otra tabla?
- A** SELECT
 - B** WHERE
 - C** UPDATE
 - D** INSERT
- 16.** ¿Cuál cláusula es requerida en una instrucción DELETE?
- 17.** ¿Cuál instrucción o cláusula se utiliza en una instrucción DELETE para especificar las filas que serán eliminadas en una tabla?

Capítulo 9

Utilizar predicados

Habilidades y conceptos clave

- Comparar datos SQL
 - Arrojar valores nulos
 - Arrojar valores similares
 - Hacer referencia a fuentes adicionales de datos
 - Determinar la cantidad de predicados de comparación
-

Hasta este punto en el libro se ha presentado gran cantidad de información acerca de diversos aspectos de los objetos de las bases de datos y los datos que éstos almacenan. En relación a esto, se analizó la consulta de datos (capítulo 7) y la modificación de datos (capítulo 8). Ahora es necesario dar un paso atrás y concentrarse en uno de los aspectos de estas discusiones: la cláusula `WHERE`. Ésta, como puede recordarse, permite especificar una condición de búsqueda que filtra aquellas filas que no se quiere que sean arrojadas por una instrucción `SELECT` o que sean modificadas por una instrucción `UPDATE` o `DELETE`. La condición de búsqueda incluye uno o más predicados que declaran cada uno un hecho acerca de la fila que será arrojada o modificada. SQL soporta varios tipos de predicados, y todos ellos permiten probar si una condición es verdadera, falsa o desconocida. Este capítulo se concentrará en aquellos predicados que son más comúnmente utilizados por programadores SQL, y se proporcionarán ejemplos sobre cómo son utilizados para ver y modificar los datos en una base de datos SQL.

Comparar datos SQL

El primer tipo de predicado que se planea analizar es el tipo de aquellos que comparan datos. Estos predicados, como cualquier predicado, están incluidos en la cláusula `WHERE`. Se puede incluir una cláusula `WHERE` en una instrucción `SELECT`, `UPDATE` o `DELETE`, y en cada caso la cláusula puede contener uno o más predicados de comparación.

Cada predicado en la cláusula `WHERE` (ya sea un predicado de comparación o de cualquier otro tipo) se evalúa en un principio individual para determinar si éste cumple la condición definida por ese predicado. Después de que se evalúan los predicados, la cláusula `WHERE` se evalúa como un todo. La cláusula debe evaluarse como verdadera si se requiere que una fila sea incluida en los resultados de una búsqueda, sea actualizada o eliminada. Si la cláusula se evaluara como falsa o desconocida, la fila no sería incluida o modificada. Para una discusión completa sobre cómo se evalúan los predicados y la cláusula `WHERE` véase el capítulo 7.

Un predicado de comparación es un tipo de predicado que compara los valores en una columna especificada con un valor especificado. Un operador de comparación se utiliza para comparar esos valores. Hemos visto ya un número de operadores de comparación (y subsecuentemente, predicados de comparación) a lo largo de este libro. La tabla 9-1 lista los seis operadores de comparación soportados por SQL y proporciona un ejemplo para cada uno.

Operador	Símbolo	Ejemplo
Igual a	=	EN_EXISTENCIA = 47
Desigual a	<>	EN_EXISTENCIA <> 47
Menor que	<	EN_EXISTENCIA < 47
Mayor que	>	EN_EXISTENCIA > 47
Menor que o igual a	<=	EN_EXISTENCIA <= 47
Mayor que o igual a	>=	EN_EXISTENCIA >= 47

Tabla 9-1 Operadores de comparación de SQL.

Sin duda usted reconocerá muchos de estos operadores, e incluso aquellos que no reconozca serán muy fáciles de comprender. Pero demos un vistazo a los ejemplos en la tabla 9-1 para comprender más cómo funciona un predicado de comparación. En la primera fila en la tabla (la fila Igual a), el predicado de ejemplo es `EN_EXISTENCIA = 47`. Si ésta fuera a aparecer en una cláusula `WHERE`, se vería de la manera siguiente:

```
WHERE EN_EXISTENCIA = 47
```

`EN_EXISTENCIA` es el nombre de la columna en la tabla identificada en la instrucción que contiene la cláusula `WHERE`. El signo de igual (=) es el operador de comparación que se utiliza para comparar los valores en la columna `EN_EXISTENCIA` con el valor a la derecha del signo de igual, el cual en este caso es 47. Por lo tanto, para que una fila sea evaluada como verdadera, el valor `EN_EXISTENCIA` para esa fila debe ser de 47. Todos los seis operadores de comparación funcionan de la misma manera. En cada caso, la cláusula `WHERE` debe evaluarse como verdadera para que la fila sea arrojada en los resultados de la consulta o para que sea modificada.

Aun cuando es tradicional colocar el nombre de la columna a la izquierda del operador de comparación y el valor de la constante a la derecha, pueden ser invertidos y formar una instrucción equivalente, asumiendo que también se ajuste el operador de comparación si fuera necesario. Por ejemplo, las siguientes dos cláusulas `WHERE` son lógicamente idénticas, y cada una selecciona filas con valores `EN_EXISTENCIA` mayores a 5:

```
WHERE EN_EXISTENCIA > 5
WHERE 5 < EN_EXISTENCIA
```

NOTA

Como se aprendió en el capítulo 7, es posible combinar predicados utilizando la palabra clave `AND` o la palabra clave `OR` para colocar juntos dos o más predicados en una cláusula `WHERE`. También se puede utilizar la palabra clave `NOT` para crear una condición inversa para un predicado particular. Recuerde, sin importar cuántos predicados sean incluidos en la cláusula `WHERE`, la cláusula aún debe evaluarse como verdadera para que una fila dada sea seleccionada.

Ahora que se ha dado un vistazo a los seis tipos de predicados de comparación, pongamos algunos ejemplos. Estos ejemplos se basan en la figura 9-1, que muestra los datos almacenados en la tabla `CDS_A_LA_MANO`.

TITULO_CD: VARCHAR(60)	DERECHOSDEAUTOR: INT	PRECIO_MENUDEO: NUMERIC(5,2)	INVENTARIO: INT
Famous Blue Raincoat	1991	16.99	6
Blue	1971	14.99	26
Court and Spark	1974	14.99	18
Past Light	1983	15.99	2
Kojiki	1990	15.99	5
That Christmas Feeling	1993	10.99	3
Patsy Cline: 12 Greatest Hits	1988	16.99	25

Figura 9-1 Comparación de datos en la tabla CDS_A_LA_MANO.

En el primer ejemplo que veremos, la cláusula WHERE utiliza un operador Igual a para comparar los valores en la columna TITULO_CD con uno de los títulos de CD:

```
SELECT TITULO_CD, DERECHOSDEAUTOR
FROM CDS_A_LA_MANO
WHERE TITULO_CD = 'Past Light';
```

Esta instrucción arrojará una fila con sólo dos valores (uno por cada columna especificada en la cláusula SELECT), como se muestra en los siguientes resultados de consulta:

```
TITULO_CD    DERECHOSDEAUTOR
-----
Past Light    1983
```

Ahora cambiemos un poco esta instrucción SELECT. En lugar de utilizar el operador Igual a, utilizaremos el operador Desigual a:

```
SELECT TITULO_CD, DERECHOSDEAUTOR
FROM CDS_A_LA_MANO
WHERE TITULO_CD <> 'Past Light';
```

Cuando se ejecuta esta instrucción, se arrojan seis filas:

```
TITULO_CD    DERECHOSDEAUTOR
-----
Famous Blue Raincoat    1991
Blue                    1971
Court and Spark          1974
Kojiki                   1990
That Christmas Feeling    1993
Patsy Cline: 12 Greatest Hits    1988
```


Observe que en los resultados de la consulta se incluyen todas las filas excepto la fila Past Light. En este caso, la cláusula WHERE se evalúa como verdadera solamente cuando el valor TITULO_CD es diferente de Past Light. Sin embargo, obsérvese que si una fila tuviera un valor nulo para TITULO_CD, esa fila no aparecería en ninguno de los conjuntos de resultados anteriores. Tenga en mente que una comparación con un valor nulo *siempre* se evaluará como desconocida (NULL) (el DBMS no puede saber si el valor es Igual a o Desigual a debido a que es un valor desconocido). Como se verá posteriormente en la sección “Arrojar valores nulos”, existen operadores especiales utilizados para tratar con los valores nulos.

Demos ahora un vistazo al operador Menor que y al operador Mayor que. En el siguiente ejemplo se combinan dos predicados de comparación utilizando la palabra clave AND:

```
SELECT TITULO_CD, INVENTARIO
FROM CDS_A_LA_MANO
WHERE INVENTARIO > 2
AND INVENTARIO < 25;
```

Como se puede ver, las filas arrojadas por esta instrucción SELECT deben contener un valor INVENTARIO entre 2 y 25. Si se ejecuta esta instrucción, se obtendrán cuatro filas:

TITULO_CD	INVENTARIO
-----	-----
Famous Blue Raincoat	6
Court and Spark	18
Kojiki	5
That Christmas Feeling	3

Al definir los predicados en una cláusula WHERE, no se está limitado a utilizar solamente una columna. Por ejemplo, si se desean seleccionar filas basadas tanto en los valores de la columna INVENTARIO y de la columna PRECIO_MENUDEO, se puede modificar la última instrucción SELECT como se muestra el siguiente ejemplo:

```
SELECT TITULO_CD, INVENTARIO
FROM CDS_A_LA_MANO
WHERE INVENTARIO > 2
AND INVENTARIO < 25
AND PRECIO_MENUDEO <> 16.99;
```

Debido a que las condiciones se conectan utilizando AND, cualquier fila que se arroje en los resultados de la consulta debe cumplir con las tres condiciones definidas en la cláusula WHERE. Como resultado, solamente se obtienen tres filas cuando se ejecuta esta instrucción:

TITULO_CD	INVENTARIO
-----	-----
Court and Spark	18
Kojiki	5
That Christmas Feeling	3

Observe que los resultados de la consulta no incluyen a la columna PRECIO_MENUDEO. Esto se debe a que esta columna no está especificada en la cláusula SELECT. Aun así, se puede utilizar esa columna en un predicado en la cláusula WHERE para definir una condición de búsqueda.

Ahora demos un vistazo al operador Menor que o igual a y al operador Mayor que o igual a. En el siguiente ejemplo, ambos operadores se utilizan para limitar las filas resultantes a aquellas con un valor DERECHOSDEAUTOR con un rango entre 1971 y 1989.

```
SELECT TITULO_CD, DERECHOSDEAUTOR
FROM CDS_A_LA_MANO
WHERE DERECHOSDEAUTOR >= 1971
      AND DERECHOSDEAUTOR <= 1989;
```

Esta instrucción arrojará resultados ligeramente diferentes a aquellos que se habrían obtenido si simplemente se hubieran utilizado los operadores Mayor que y Menor que. Utilizando el operador Mayor que o igual a y el operador Menor que o igual a, los valores iguales al valor especificado también son arrojados, como se muestra en los siguientes resultados de consulta:

TITULO_CD	DERECHOSDEAUTOR
-----	-----
Blue	1971
Court and Spark	1974
Past Light	1983
Patsy Cline: 12 Greatest Hits	1988

Observe que la fila Blue incluye un valor DERECHOSDEAUTOR de 1971. Éste no habría sido incluido si se hubiera utilizado solamente el operador Mayor que.

Hasta este punto, todos los ejemplos mostrados han estado basados en las instrucciones SELECT. Sin embargo, se puede agregar una cláusula WHERE a una instrucción UPDATE o a una instrucción DELETE. Supongamos que se quiere incrementar el valor INVENTARIO para la fila That Christmas Feeling. Se puede utilizar la siguiente instrucción UPDATE:

```
UPDATE CDS_A_LA_MANO
      SET INVENTARIO = 10
      WHERE TITULO_CD = 'That Christmas Feeling';
```

Cuando se ejecuta esta instrucción, el valor INVENTARIO es incrementado a 10 para la fila That Christmas Feeling, pero únicamente para esa fila debido a que la cláusula WHERE se evalúa como verdadera sólo para esa fila. De manera muy sencilla se pudo haber agregado esta cláusula WHERE a una instrucción DELETE, en cuyo caso la fila That Christmas Feeling hubiera sido eliminada.

Al igual que con la cláusula WHERE en una instrucción SELECT, es posible combinar dos o más predicados para formar una condición de búsqueda:

```
UPDATE CDS_A_LA_MANO
      SET INVENTARIO = 3
      WHERE TITULO_CD = 'That Christmas Feeling'
            AND DERECHOSDEAUTOR = 1993;
```

Cuando se especifica la palabra clave AND, ambos predicados deben evaluarse como verdaderos para que la cláusula WHERE pueda evaluarse como verdadera. Si se especifica la palabra clave OR en lugar de AND, entonces sólo será necesario que un predicado se evalúe como verdadero para que la fila sea seleccionada.

Utilizar el predicado BETWEEN

Hablando estrictamente, el predicado BETWEEN no es un predicado de comparación, al menos no lo es tal como se presenta en SQL:2006 estándar. Sin embargo, es lo suficientemente similar en funciones al operador Mayor que o Igual a y a los operadores Menor que o Igual a, que vale la pena analizar aquí.

El predicado BETWEEN se utiliza en conjunción con la palabra clave AND para identificar un rango de valores que pueden ser incluidos como una condición de búsqueda en la cláusula WHERE. Los valores en la columna identificada deben entrar en ese rango para poder evaluarse como verdaderos. Cuando se utiliza la cláusula BETWEEN, se debe especificar la columna aplicable, el valor más bajo del rango y el valor más alto del rango. El siguiente ejemplo (que está basado en la tabla CDS_A_LA_MANO en la figura 9-1) especifica un rango entre 14 y 16.

```
SELECT TITULO_CD, PRECIO_MENUDEO
FROM CDS_A_LA_MANO
WHERE PRECIO_MENUDEO BETWEEN 14 AND 16;
```

El valor PRECIO_MENUDEO para cada fila seleccionada debe entrar en ese rango, incluyendo los puntos límite. Si se ejecuta esta instrucción, solamente cuatro filas serán incluidas en los resultados de la consulta:

TITULO_CD	PRECIO_MENUDEO
-----	-----
Blue	14.99
Court and Spark	14.99
Past Light	15.99
Kojiki	15.99

Ahora demos un vistazo a una consulta similar a la presentada en el último ejemplo, sólo que esta vez utilizando los predicados de comparación en lugar del predicado BETWEEN:

```
SELECT TITULO_CD, PRECIO_MENUDEO
FROM CDS_A_LA_MANO
WHERE PRECIO_MENUDEO >= 14
AND PRECIO_MENUDEO <= 16;
```

Observe que se utilizan dos predicados: uno con el operador Mayor que o igual a, y el otro con el operador Menor que o igual a. Esta instrucción SELECT producirá los mismos resultados de búsqueda que la instrucción SELECT anterior.

Ahora regresemos al predicado BETWEEN. Al igual que con cualquier predicado, se puede combinar el predicado BETWEEN con otros predicados. En la siguiente instrucción, la cláusula WHERE incluye un predicado BETWEEN y un predicado de comparación:

```
SELECT TITULO_CD, PRECIO_MENUDEO
FROM CDS_A_LA_MANO
WHERE PRECIO_MENUDEO BETWEEN 14 AND 16
AND INVENTARIO > 10;
```

Como resultado de ambos predicados, los resultados de la consulta pueden incluir solamente aquellas filas con un valor `PRECIO_MENUDEO` que quepa en el rango entre 14 y 16 y con un valor `INVENTARIO` mayor a 10. Cuando se ejecuta esta consulta, solamente se obtienen dos filas:

TITULO_CD	PRECIO_MENUDEO
-----	-----
Blue	14.99
Court and Spark	14.99

Una vez más, puede notarse que los resultados de la consulta no incluyen a la columna `INVENTARIO` incluso cuando esa columna se especifica en un predicado dentro de la cláusula `WHERE`. Observe también que se hace referencia a más de una columna en la cláusula `WHERE`.

Además de lo que se ha visto hasta ahora para el predicado `BETWEEN`, también puede utilizarse la cláusula para especificar el inverso de una condición. Esto se hace utilizando la palabra clave `NOT` dentro del predicado. Por ejemplo, supongamos que se cambia el último ejemplo a lo siguiente:

```
SELECT TITULO_CD, PRECIO_MENUDEO
FROM CDS_A_LA_MANO
WHERE PRECIO_MENUDEO NOT BETWEEN 14 AND 16;
```

Las filas arrojadas en los resultados de la consulta incluirán todas las filas que no tengan un valor `PRECIO_MENUDEO` dentro del rango entre 14 y 16. Cuando se ejecuta la instrucción, se obtienen tres filas:

TITULO_CD	PRECIO_MENUDEO
-----	-----
Famous Blue Raincoat	16.99
That Christmas Feeling	10.99
Patsy Cline: 12 Greatest Hits	16.99

Observe que todos los valores dentro del rango especificado han sido excluidos de los resultados de la consulta. Si usted encuentra confusa la función de la palabra clave `NOT` (como le pasa a mucha gente), puede escribir una cláusula equivalente `WHERE` o utilizar la palabra clave `OR` y los operadores Mayor que y Menor que de la manera siguiente:

```
WHERE PRECIO_MENUDEO < 14
      OR PRECIO_MENUDEO > 16;
```

Arrojar valores nulos

Como se puede recordar del capítulo 4, un valor nulo es utilizado en lugar de un valor cuando dicho valor no está definido o no es conocido. Un valor nulo indica que el valor está ausente. Esto no es lo mismo que un valor cero, un espacio en blanco o un valor por defecto. Por opción predeterminada, SQL permite utilizar valores nulos en lugar de valores regulares (aunque se puede anular el valor por opción predeterminada incluyendo una restricción `NOT NULL` en la definición de la columna). En aquellos casos donde se permiten los valores nulos, puede ser necesario especificar que esos valores nulos sean arrojados cuando se consulta una tabla. Por esta razón, SQL proporciona un predicado `NULL`, que permite definir las condiciones de búsqueda que arrojaran los valores nulos.

El predicado NULL resulta muy simple de implementar. Utilizado en conjunción con la palabra clave IS, el predicado se agrega a una cláusula WHERE de la misma forma que cualquier otro predicado, y se aplica sólo a los valores nulos que pudieran existir en la columna que se está consultando. La mejor forma de ilustrar esto es mediante el uso de ejemplos. En estos ejemplos se utiliza la tabla BIO_ARTISTAS, mostrada en la figura 9-2.

El primer ejemplo es una instrucción SELECT que arroja filas con un valor LUGAR_DE_NACIMIENTO nulo:

```
SELECT *
FROM BIO_ARTISTAS
WHERE LUGAR_DE_NACIMIENTO IS NULL;
```

La instrucción arroja todas las columnas de la tabla BIO_ARTISTAS; sin embargo, arroja solamente dos filas, como se puede ver en los siguientes resultados de la consulta (los resultados pueden lucir diferentes dependiendo de cómo la aplicación cliente de SQL despliega los valores nulos):

NOMBRE_INTERPRETE	LUGAR_DE_NACIMIENTO	AÑO_NACIMIENTO
William Ackerman	NULL	NULL
Bing Crosby	NULL	1904

El hecho de que la columna AÑO_NACIMIENTO contenga un valor nulo para la fila William Ackerman no tiene relación directa con el hecho de que se utilice un predicado NULL. En este caso el predicado NULL solamente identifica a la columna LUGAR_DE_NACIMIENTO, no a la columna AÑO_NACIMIENTO. Sin embargo, es posible reemplazar la columna LUGAR_DE_NACIMIENTO en el predicado con la columna AÑO_NACIMIENTO, en cuyo caso las filas arrojadas serán aquellas con un valor AÑO_NACIMIENTO nulo.

NOMBRE_INTERPRETE: VARCHAR(60)	LUGAR_DE_NACIMIENTO CARCHAR(60)	AÑO_NACIMIENTO: INT
Jennifer Warnes	Seattle, Washington, USA	1947
Joni Mitchell	Fort MacLeod, Alberta, Canada	1943
William Ackerman	NULL	NULL
Kitaro	Toyohashi, Japan	NULL
Bing Crosby	NULL	1904
Patsy Cline	Winchester, Virginia, United States	1932
Jose Carreras	Barcelona, Spain	NULL
Luciano Pavarotti	Modena, Italy	1935
Placido Domingo	Madrid, Spain	1941

Figura 9-2 Valores nulos arrojados de la tabla BIO_ARTISTAS.

De acuerdo al estándar SQL:2006, también es posible especificar ambas columnas en el predicado NULL, como se muestra en el siguiente ejemplo:

```
SELECT *
  FROM BIO_ARTISTAS
 WHERE (LUGAR_DE_NACIMIENTO, AÑO_NACIMIENTO) IS NULL;
```

Cuando se incluyen ambas columnas, tanto la columna LUGAR_DE_NACIMIENTO como la columna AÑO_NACIMIENTO deben arrojar valores nulos para que una fila aparezca en la búsqueda, que en el caso de la tabla BIO_ARTISTAS sería solamente una fila.

NOTA

A pesar de que el estándar SQL permite especificar múltiples columnas en el predicador NULL, existen muchas implementaciones que no pueden soportar esto. En su lugar, se deben especificar dos predicados NULL conectados con la palabra clave AND.

Como una alternativa para incluir ambas columnas en un predicado, se puede escribir la instrucción SELECT de la forma siguiente:

```
SELECT *
  FROM BIO_ARTISTAS
 WHERE LUGAR_DE_NACIMIENTO IS NULL
        AND AÑO_NACIMIENTO IS NULL;
```

Si se ejecuta esta instrucción, se obtendrán los siguientes resultados de consulta:

PERFORMER_NAME	LUGAR_DE_NACIMIENTO	AÑO_NACIMIENTO
-----	-----	-----
William Ackerman	NULL	NULL

SQL soporta otra característica en el predicador NULL. Se puede utilizar la palabra clave NOT para encontrar los resultados inversos de predicado. Por ejemplo, supongamos que se quieren obtener todas las filas que incluyan un valor actual en la columna LUGAR_DE_NACIMIENTO en lugar de un valor nulo. La instrucción podría lucir de esta manera:

```
SELECT *
  FROM BIO_ARTISTAS
 WHERE LUGAR_DE_NACIMIENTO IS NOT NULL;
```

Los resultados de la consulta ahora incluirán siete filas, todas ellas conteniendo valores en la columna LUGAR_DE_NACIMIENTO:

NOMBRE_INTERPRETE	LUGAR_DE_NACIMIENTO	AÑO_NACIMIENTO
-----	-----	-----
Jennifer Warnes	Seattle, Washington, Estados Unidos	1947
Joni Mitchell	Fort MacLeod, Alberta, Canadá	1943
Kitaro	Toyohashi, Japón	NULL
Patsy Cline	Winchester, Virginia, Estados Unidos	1932
Jose Carreras	Barcelona, España	NULL

Luciano Pavarotti	Modena, Italia	1935
Plácido Domingo	Madrid, España	1941

Observe que los valores nulos pueden aún existir en otras columnas. Debido a que solamente la columna LUGAR_DE_NACIMIENTO está especificada en el predicado NULL, solamente esa columna debe contener un valor para que una fila sea arrojada.

Al igual que con los predicados que se vieron anteriormente en este capítulo, se puede combinar el predicado NULL con otros tipos de predicados. Por ejemplo, se puede modificar el último ejemplo para limitar los valores de AÑO_NACIMIENTO a sólo ciertos años, como se muestra en el siguiente ejemplo:

```
SELECT *
FROM BIO_ARTISTAS
WHERE LUGAR_DE_NACIMIENTO IS NOT NULL
AND AÑO_NACIMIENTO > 1940;
```

Ahora cualquier fila arrojada deberá incluir un valor en la columna LUGAR_DE_NACIMIENTO y el valor de AÑO_NACIMIENTO deberá ser mayor a 1940. Si se ejecuta esta consulta, se arrojarán los siguientes resultados:

NOMBRE_INTERPRETE	LUGAR_DE_NACIMIENTO	AÑO_NACIMIENTO
Jennifer Warnes	Seattle, Washington, Estados Unidos	1947
Joni Mitchell	Fort MacLeod, Alberta, Canadá	1943
Plácido Domingo	Madrid, España	1941

Como se puede ver, solamente se obtienen tres filas. Ninguna fila con valor LUGAR_DE_NACIMIENTO nulo es arrojada debido a que el valor nulo se evalúa como desconocido, y solamente las cláusulas WHERE que se evalúan como verdaderas pueden ser incluidas en los resultados de la consulta.

Arrojar valores similares

Si algún predicado puede ser divertido, éste es el predicado LIKE. El predicado LIKE proporciona un ambiente flexible en el cual es posible especificar valores que son solamente *similares* a los valores almacenados en la base de datos. Esto es particularmente beneficioso si sólo se conoce parte de un valor pero aún se necesita recuperar información basada en ese valor. Por ejemplo, supongamos que no se conoce el título completo de un CD, sino solamente una parte de ese título. O quizá sólo se conoce una parte del nombre del artista. Al utilizar el predicado LIKE, se pueden solicitar valores que sean similares a la parte que se conoce y desde esos resultados determinar si la información que se necesita está ahí.

Antes de dar un vistazo al predicado LIKE en sí mismo, veamos dos símbolos utilizados dentro del predicado. El predicado LIKE utiliza dos caracteres especiales, el signo de porcentaje (%) y el guión bajo (_), para ayudar a definir las condiciones de búsqueda especificadas en el predicado. El signo de porcentaje representa cero o más caracteres desconocidos, y el guión bajo representa exactamente un carácter desconocido. Se pueden utilizar estos caracteres al inicio de un valor, en medio de él, o al final, y se pueden combinar entre sí según sea necesario. La forma en que se utilizan estos dos caracteres determina el tipo de datos que serán recuperados de la base de datos.

Valor de ejemplo	Posibles resultados de la consulta
'J%'	Jennifer Warnes, Joni Mitchell, Jose Carreras
'%Spark'	Court and Spark
'%Blue%'	Famous Blue Raincoat, Blue, Blues on the Bayou
'%Cline%Hits'	Patsy cline: 12 Greatest Hits
'194_'	1940, 1941, 1947
'19__'	1900, 1907, 1938, 1963, 1999
'__ue'	Blue
'9__01'	90201, 91401, 95501, 99301, 99901
'9_3%'	9032343, 903, 95312, 99306, 983393300333

Tabla 9-2 Utilizar caracteres especiales en un predicado LIKE.

La tabla 9-2 proporciona varios ejemplos de cómo pueden ser utilizados estos caracteres especiales en un predicado LIKE.

Como se puede ver, los caracteres de signo de porcentaje y el guión bajo proporcionan una gran flexibilidad y permiten consultar un amplio rango de datos.

NOTA

Algunas implementaciones de los fabricantes pueden ser configuradas para ser insensibles al uso de mayúsculas y minúsculas para las comparaciones de datos, lo que significa que las letras en minúsculas o en mayúsculas en una comparación se consideran iguales. De hecho, éste es el comportamiento por defecto para SQL Server y Sybase; sin embargo, Oracle siempre es sensible a mayúsculas y minúsculas. Observando la tabla 9-2, las letras en mayúscula en la columna Valor de ejemplo pueden ser cambiadas a minúsculas con los mismos resultados siempre que DBMS esté configurado como insensible al uso de mayúsculas y minúsculas. Por ejemplo, el valor %spark del predicado LIKE aún seleccionaría a la fila Court and Spark en una implementación insensible al uso de mayúsculas y minúsculas.

Ahora que están mejor comprendidos los caracteres especiales, demos un vistazo al predicado LIKE como un todo. El predicado LIKE incluye el nombre de la columna, la palabra clave LIKE, y un valor encerrado en un conjunto de comillas simples, a su vez encerrado en un conjunto de paréntesis (los paréntesis son opcionales para la mayoría de las implementaciones de los fabricantes). Por ejemplo, la siguiente cláusula WHERE incluye un predicado LIKE:

```
WHERE ID_CD LIKE ('%01')
```

El predicado incluye la columna ID_CD, la palabra clave LIKE y un valor de %01. Sólo las filas que contengan el valor correcto en la columna ID_CD son arrojadas en los resultados de la consulta. La columna ID_CD es parte de la tabla CDS, que se muestra en la figura 9-3. Se estará utilizando esta tabla para los ejemplos en esta sección. Observe que basándose en el predicado LIKE definido en la cláusula WHERE precedente, solamente una fila puede ser arrojada por esta cláusula, la fila con el valor ID_CD de 99301.

ID_CD: INT	TITULO_CD: VARCHAR(60)
99301	Famous Blue Raincoat
99302	Blue
99303	Court and Spark
99304	Past Light
99305	Kojiki
99306	That Christmas Feeling
99307	Patsy Cline: 12 Greatest Hits

Figura 9-3 Arrojar valores similares desde la tabla CDS.

Ahora demos un vistazo a algunos ejemplos de instrucciones SELECT que incluyan un predicado LIKE. Supongamos que se quiere encontrar cualquier CD que contenga la palabra Christmas en el título. Se puede crear la siguiente instrucción SELECT para consultar la tabla CDS:

```
SELECT *
  FROM CDS
 WHERE TITULO_CD LIKE ('%Christmas%');
```

Los resultados de esta consulta incluirán solamente una fila:

```
ID_CD      TITULO_CD
-----
99306      That Christmas Feeling
```

Si se hubiera incluido solamente un signo de porcentaje, no se habría obtenido ninguna fila. Por ejemplo, si se eliminara el primer signo de porcentaje, la implementación SQL habría interpretado esto como el significado de que el valor debe comenzar con la palabra Christmas, lo cual no es así. Lo mismo sucede para el otro signo de porcentaje. Si éste se hubiera eliminado, la implementación habría asumido que Christmas debe ser la última palabra de la cadena de caracteres. Además de esto, si no se hubiera utilizado ningún signo de porcentaje, ninguna fila hubiera sido arrojada porque ningún valor habría coincidido exactamente con la palabra Christmas.

También se puede agregar la palabra clave NOT a un predicado LIKE si quieren obtenerse todas las filas, excepto aquellas especificadas por el predicado. Tomemos el último ejemplo. Si se agregara la palabra clave NOT, luciría de la siguiente manera:

```
SELECT *
  FROM CDS
 WHERE TITULO_CD NOT LIKE ('%Christmas%');
```

Esta vez los resultados de la consulta incluyen todas las filas que no contengan la palabra Christmas:

ID_CD	TITULO_CD
-----	-----
99301	Famous Blue Raincoat
99302	Blue
99303	Court and Spark
99304	Past Light
99305	Kojiki
99307	Patsy Cline: 12 Greatest Hits

Observe que ahora falta la fila That Christmas Feeling.

También se puede combinar un predicado LIKE con otro predicado LIKE. Supongamos, por ejemplo, que aún se quiere excluir el valor Christmas, pero se quiere incluir el valor Blue, como se muestra en el siguiente ejemplo:

```
SELECT *
FROM CDS
WHERE TITULO_CD NOT LIKE ('%Christmas%')
AND TITULO_CD LIKE ('%Blue%');
```

La cláusula WHERE en la instrucción SELECT elimina cualquier fila que tenga la palabra Christmas en cualquier parte del valor TITULO_CD. Además, el valor TITULO_CD debe incluir la palabra Blue. Como resultado, sólo se obtienen dos filas.

ID_CD	TITULO_CD
-----	-----
99301	Famous Blue Raincoat
99302	Blue

Pero ¿qué sucede si el título del CD incluye ambas palabras? Por ejemplo, *Blue Christmas* de Elvis Presley que está ahora disponible en CD. La palabra clave AND utilizada para conectar los predicados significa que *ambos* predicados deben ser verdaderos para que una fila pueda ser arrojada. Incluso si existía una fila Blue Christmas, no sería incluida en los resultados de la consulta debido a que el primer predicado (el que tiene NOT LIKE) sería evaluado como falso.

Pruebe esto 9-1

Utilizar predicados en instrucciones SQL

Antes de moverse hacia otros predicados, es una buena idea revisar los predicados que ya han sido discutidos. Éstos incluyen los seis tipos de predicados de comparación, el predicado BETWEEN, el predicado NULL y el predicado LIKE. En este ejercicio se probarán varios de estos predicados mediante el uso de instrucciones SELECT que incluirán las cláusulas WHERE apropiadas. Se consultarán las tablas que se crearon en la base de datos INVENTARIO. Debido a que se utilizarán solamente instrucciones SELECT, de cualquier manera no se modificarán las tablas o las estructuras de las bases de datos. Solamente se solicitarán datos basándose en los predicados que se definen. Puede descargar el archivo Try_This_09.txt (en inglés), que contiene las instrucciones SQL utilizadas en este ejercicio.

Paso a paso

1. Abra la aplicación de cliente para sus RDBMS y conéctese con la base de datos INVENTARIO.
2. En la primera instrucción creada, se consultará la tabla TIPOS_MUSICA para arrojar los nombres de aquellas filas cuyo valor ID_TIPO sea igual a 11 o 12. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT ID_TIPO, NOMBRE_TIPO
FROM TIPOS_MUSICA
WHERE ID_TIPO = 11
      OR ID_TIPO = 12;
```

La instrucción deberá arrojar dos filas, una para Blues y otra para Jazz. Observe que la palabra clave OR se utiliza para indicar que cualesquiera de los dos valores es aceptable.

3. Ahora se consultará la tabla ARTISTAS para buscar artistas diferentes a Patsy Cline y Bing Crosby. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT NOMBRE_ARTISTA, LUGAR_DE_NACIMIENTO
FROM ARTISTAS
WHERE NOMBRE_ARTISTA <> 'Patsy Cline'
      AND NOMBRE_ARTISTA <> 'Bing Crosby';
```

La consulta deberá arrojar 16 filas y no deberá incluir las filas Patsy Cline o Bing Crosby.

4. Ahora combinemos un par de predicados de comparación para crear un tipo diferente de condición de búsqueda. En esta instrucción se consultará una vez más la tabla ARTISTAS, pero ahora sólo se petitionarán aquellas filas cuyos valores ID_ARTISTA se encuentren entre 2004 y 2014 (excluyendo los puntos límite). Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT ID_ARTISTA, NOMBRE_ARTISTA
FROM ARTISTAS
WHERE ID_ARTISTA > 2004
      AND ID_ARTISTA < 2014;
```

La consulta deberá arrojar nueve filas.

5. Ahora modifiquemos la instrucción SELECT que se acaba de ejecutar. Deberá utilizarse un predicado BETWEEN en lugar de los dos predicados de comparación. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT ID_ARTISTA, NOMBRE_ARTISTA
FROM ARTISTAS
WHERE ID_ARTISTA BETWEEN 2004 AND 2014;
```

Ahora deberán aparecer 11 filas, en lugar de las nueve que se arrojaron en el paso anterior, debido a que BETWEEN siempre incluye los puntos límite. Si en el paso anterior se hubiera

(continúa)

utilizado el operador Mayor que o igual a y el operador Menor que o igual a, los resultados de la consulta habrían sido los mismos que en este paso.

- 6.** Ahora consultemos una vez más la tabla ARTISTAS, sólo que esta vez se utilizará el predicado NULL. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT *
  FROM ARTISTAS
 WHERE LUGAR_DE_NACIMIENTO IS NULL;
```

La búsqueda no arrojará resultados debido a que la columna LUGAR_DE_NACIMIENTO no contiene valores nulos.

- 7.** Intentemos la misma consulta que en el paso anterior, sólo que esta vez se agregará la palabra clave NOT al predicado NULL. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT *
  FROM ARTISTAS
 WHERE LUGAR_DE_NACIMIENTO IS NOT NULL;
```

La consulta ahora deberá arrojar todas las filas de la tabla (18 en total).

- 8.** En la siguiente instrucción se utilizará el predicado LIKE para encontrar títulos de CD que incluyan la palabra Best o la palabra Greatest. El predicado hará referencia a la columna TITULO_CD de la tabla DISCOS_COMPACTOS. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT TITULO_CD, EN_EXISTENCIA
  FROM DISCOS_COMPACTOS
 WHERE TITULO_CD LIKE ('%Greatest%')
        OR TITULO_CD LIKE ('%Best%');
```

La consulta deberá arrojar tres filas. En todas estas filas, el valor TITULO_CD deberá contener las palabras Greatest o Best.

- 9.** Esta vez se modificará la instrucción en el paso anterior para incluir la palabra clave NOT en ambos predicados. También deberá cambiarse la palabra clave OR por la palabra AND. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT TITULO_CD, EN_EXISTENCIA
  FROM DISCOS_COMPACTOS
 WHERE TITULO_CD NOT LIKE ('%Greatest%')
        AND TITULO_CD NOT LIKE ('%Best%');
```

Los resultados de la consulta deberán ahora incluir 12 filas. Si no se hubiera cambiado la palabra clave OR por AND, los resultados habrían incluido todas las 15 filas. Esto se debe a que la instrucción *siempre* se habría evaluado como verdadera (el primer predicado se evaluaría como verdadero para una fila que contenga Greatest; el segundo predicado se evaluaría como verdadero para una fila que contenga Best, y ambos predicados se evaluarían como verdaderos para cualquier otra fila).

- 10.** Cierre la aplicación de cliente.

Resumen de Pruebe esto

En este ejercicio se crearon varias instrucciones `SELECT` que incluían varios predicados. Los predicados estaban contenidos en cláusulas `WHERE` que eran parte de las instrucciones `SELECT`; sin embargo, estas cláusulas también pudieron haber sido parte de las instrucciones `UPDATE` y `DELETE`. Al avanzar en el resto de este capítulo, se aprenderá acerca de otros predicados y cómo pueden ser utilizados en diferentes tipos de instrucciones. Estos predicados pueden ser utilizados en conjunción con los que ya han sido analizados utilizándolos por sí mismos para crear condiciones de búsqueda más complejas y para arrojar resultados más precisos.

Hacer referencia a fuentes adicionales de datos

SQL soporta muchos tipos de predicados que permiten hacer referencia a fuentes diferentes a la tabla principal que se está consultando o modificando. Como resultado, se pueden crear condiciones de búsqueda que comparen datos entre tablas para determinar cuáles filas deberán ser incluidas en los resultados de la consulta, cuáles filas deberán ser actualizadas o cuáles eliminadas. En esta sección se verán dos predicados importantes que se pueden utilizar para hacer referencia a otras tablas: el predicado `IN` y el predicado `EXISTS`.

Ambos predicados pueden utilizar subconsultas para hacer referencia a datos en la tabla que se está consultando o modificando, o más comúnmente, a otras tablas. Se introdujo por primera vez el tema de las subconsultas en el capítulo 4. Como se puede recordar de ese capítulo, una subconsulta es una expresión que se utiliza como un componente dentro de otra expresión. En su uso más común, una subconsulta es simplemente una instrucción `SELECT` incrustada dentro de otra instrucción. Cuando se utiliza en un predicado, una subconsulta se convierte en parte de ese predicado y consecuentemente está incrustada en la cláusula `WHERE` de una instrucción `SELECT`, `UPDATE` o `DELETE`. A pesar de que las subconsultas se analizan con detalle en el capítulo 12, se mencionan aquí debido a que son una parte integral de los predicados que se estarán discutiendo en la siguiente parte de este capítulo. En cada uno de estos predicados, las subconsultas se utilizan para hacer referencia a datos en otras tablas. Para los propósitos de este capítulo, se mantendrán los ejemplos y las subconsultas simples, pero tenga en mente que éstas pueden ser mucho más elaboradas que lo que se verá aquí, y una vez que haya completado el capítulo 12, será capaz de aplicar ese conocimiento a los predicados que se aprendieron en este capítulo.

Utilizar el predicado `IN`

El predicado `IN` permite determinar si los valores en la columna especificada de una tabla están contenidos en una lista definida o contenidos dentro de otra tabla. En el primer caso, se debe especificar el nombre de la columna, la palabra clave `IN` y una lista de valores que son comparados a los valores en la columna especificada. En el segundo caso, se debe especificar el nombre de la columna, la palabra clave `IN` y una subconsulta, que hace referencia a la segunda tabla. En cada caso, si el valor de la columna coincide con uno de los valores en la lista o en los resultados de la subconsulta, el predicado se evalúa como verdadero y la fila es arrojada en los resultados de la consulta.

INVENTARIO_DISCO_COMPACTO		ARTISTAS_DISCO_COMPACTO	
NOMBRE_CD: VARCHAR(60)	EN_EXISTENCIA: INT	TITULO: VARCHAR(60)	ARTISTA: VARCHAR(60)
Famous Blue Raincoat	13	Famous Blue Raincoat	Jennifer Warnes
Blue	42	Blue	Joni Mitchell
Court and Spark	22	Court and Spark	Joni Mitchell
Past Light	17	Past Light	William Ackerman
Kojiki	6	Kojiki	Kitaro
That Christmas Feeling	8	That Christmas Feeling	Bing Crosby
Out of Africa	29	Patsy Cline: 12 Greatest Hits	Patsy Cline
Blues on the Bayou	27	After the Rain: The Soft Sounds of Erik Satie	Pascal Roge
Orlando	5	Out of Africa	John Barry
		Leonard Cohen The Best of	Leonard Cohen
		Fundamental	Bonnie Raitt
		Blues on the Bayou	B.B. King
		Orlando	David Motion

Figura 9-4 Consultar datos de la tabla INVENTARIO_DISCO_COMPACTO y de la tabla ARTISTAS_DISCO_COMPACTO.

La mejor forma de ilustrar ambos de estos métodos es por medio de ejemplos. Sin embargo, antes de verlos, refiérase a las tablas mostradas en la figura 9-4. Éstas son las tablas que se estarán utilizando para los ejemplos.

Como ya se mencionó, el primer método para utilizar el predicado IN es definir una lista. La lista deberá incluir todos los valores que van a ser comparados a los valores en la columna especificada. Por ejemplo, supongamos que se quieren limitar los resultados de la consulta a las filas en la tabla INVENTARIO_DISCO_COMPACTO que tengan un valor EN_EXISTENCIA de 12, 22, 32 o 42. Se puede crear una instrucción SELECT como la siguiente:

```
SELECT NOMBRE_CD, EN_EXISTENCIA FROM INVENTARIO_DISCO_COMPACTO
WHERE EN_EXISTENCIA IN ( 12, 22, 32, 42 );
```

Esta instrucción arroja solamente dos filas debido a que son las únicas filas que tienen los valores EN_EXISTENCIA correctos:

```
NOMBRE_CD          EN_EXISTENCIA
-----
Blue               42
Court and Spark    22
```

Como se puede ver, es un proceso bastante sencillo utilizar el predicado IN para definir una lista, y es bastante útil cuando se sabe exactamente con cuáles valores se quieren comparar las columnas. También es un método más simple que definir predicados separados para cada valor, como en el siguiente ejemplo:

```
SELECT NOMBRE_CD, EN_EXISTENCIA
FROM INVENTARIO_DISCO_COMPACTO
WHERE EN_EXISTENCIA = 12
      OR EN_EXISTENCIA = 22
      OR EN_EXISTENCIA = 32
      OR EN_EXISTENCIA = 42;
```

Esta instrucción arrojará los mismos resultados que la instrucción SELECT del ejemplo anterior; sin embargo, como se puede ver, es mucho más incómoda.

Ahora veamos una instrucción SELECT que utiliza una subconsulta en el predicado IN. Supongamos que se quiere crear una consulta que arroje nombres de CD y sus artistas. Se requiere que los resultados de la consulta incluyan solamente CD que tengan más de 10 copias en existencia. Si nos referimos a la figura 9-4, se verá que la tabla ARTISTAS_DISCO_COMPACTO incluye los nombres de CD y sus artistas. Sin embargo, como también se puede ver, los valores EN_EXISTENCIA están almacenados en la tabla INVENTARIO_DISCO_COMPACTO, lo que significa que se necesitará hacer referencia a esa tabla para poder arrojar las filas correctas. Para hacer eso, se puede crear la siguiente instrucción SELECT:

```
SELECT TITULO, ARTISTA
FROM ARTISTAS_DISCO_COMPACTO
WHERE TITULO IN
      ( SELECT NOMBRE_CD
        FROM INVENTARIO_DISCO_COMPACTO
        WHERE EN_EXISTENCIA > 10 );
```

Si se ejecuta esta instrucción, se recibirán los siguientes resultados:

TITULO	ARTISTA
Famous Blue Raincoat	Jennifer Warnes
Blue	Joni Mitchell
Court and Spark	Joni Mitchell
Past Light	William Ackerman
Out of Africa	John Barry
Blues on the Bayou	B.B. King

Observe que sólo seis filas han sido arrojadas. Sólo hay seis CD enlistados en la tabla INVENTARIO_DISCO_COMPACTO que tienen un valor EN_EXISTENCIA mayor a 10.

Ahora veamos más de cerca la instrucción SELECT para obtener una mejor comprensión de cómo funciona el predicado IN. La cláusula WHERE contiene solamente un predicado. Éste comienza con el nombre de la columna (TITULO) cuyos valores se quieren verificar. La columna TITULO es seguida por la palabra clave IN. Luego la palabra clave es seguida por una sub-

consulta, que está encerrada en paréntesis. La subconsulta consiste de la siguiente instrucción SELECT:

```
SELECT NOMBRE_CD
      FROM INVENTARIO_DISCO_COMPACTO
      WHERE EN_EXISTENCIA > 10
```

Si se ejecutara sólo esta instrucción, se arrojarían los siguientes resultados:

```
NOMBRE_CD
-----
Famous Blue Raincoat
Blue
Court and Spark
Past Light
Out of Africa
Blues on the Bayou
```

Cada fila en los resultados de la consulta, que se derivan de la tabla INVENTARIO_DISCO_COMPACTO, contiene un valor EN_EXISTENCIA mayor a 10. Los valores en la columna TITULO de la tabla ARTISTAS_DISCO_COMPACTO son entonces comparados con estos resultados. Cualquier fila que contenga un valor TITULO que coincida con uno de los seis valores NOMBRE_CD (en los resultados de la subconsulta) está incluida en los resultados de la consulta de la instrucción SELECT principal.

NOTA

Cuando se incluye una subconsulta en un predicado IN, la cláusula SELECT de la subconsulta debe arrojar solamente una columna de datos. Si se especifica más de una columna en el conjunto de resultados o se especifica un asterisco, se recibirá un error.

Al igual que muchos otros predicados, el predicado IN permite especificar el inverso de una condición al utilizar la palabra clave NOT. Supongamos que se reescribe la instrucción SELECT del último ejemplo para incluir la palabra clave NOT en el predicado IN:

```
SELECT TITULO, ARTISTA
      FROM ARTISTAS_DISCO_COMPACTO
      WHERE TITULO NOT IN
        ( SELECT NOMBRE_CD
          FROM INVENTARIO_DISCO_COMPACTO
          WHERE EN_EXISTENCIA > 10 );
```

Los resultados de la consulta incluirán todas aquellas filas que no fueron arrojadas por la instrucción SELECT anterior y excluirá aquellas que sí fueron arrojadas, como se muestra en los siguientes resultados:

TITULO	ARTISTA
-----	-----
Kojiki	Kitaro
That Christmas Feeling	Bing Crosby
Patsy Cline: 12 Greatest Hits	Patsy Cline

After the Rain: The Soft Sounds of Erik Satie
 Leonard Cohen The Best of
 Fundamental
 Orlando

Pascal Roge
 Leonard Cohen
 Bonnie Raitt
 David Motion

Como se puede ver, el predicado IN es una herramienta muy flexible para comparar valores en una columna especificada con datos en otras tablas. Usted encontrará esto extremadamente útil cuando aprenda más acerca de las subconsultas y pueda crear predicados más complejos.

Utilizar el predicado EXISTS

A pesar de ser similar al predicado IN, el predicado EXISTS tiene un enfoque ligeramente diferente. Está dedicado únicamente a determinar si la subconsulta arroja alguna fila o no. Si ésta arroja una o más filas, el predicado se evalúa como verdadero; de otra manera, el predicado se evalúa como falso. El predicado consiste de la palabra clave EXISTS y una subconsulta. Para que la subconsulta sea un valor real (y subsecuentemente el predicado EXISTS por sí mismo), debe incluir un predicado que coincida con dos columnas en diferentes tablas. Por ejemplo, en la figura 9-4, la tabla INVENTARIO_DISCO_COMPACTO incluye la columna TITULO. Las dos columnas pueden hacerse coincidir juntas para asegurar que sólo las filas relevantes sean arrojadas por la subconsulta. Veamos un ejemplo para ayudar a aclarar este tema.

Supongamos que se quieren recuperar filas desde la tabla INVENTARIO_DISCO_COMPACTO para poder determinar cuántos CD de Joni Mitchell se tienen en existencia. Solamente se quieren desplegar los nombres de los CD y el número de ellos en existencia. No se quiere desplegar el nombre del artista, ni los CD de otros artistas. Para cumplir con esto, se puede utilizar la siguiente instrucción SELECT:

```
SELECT *
  FROM INVENTARIO_DISCO_COMPACTO
 WHERE EXISTS
    ( SELECT TITULO
      FROM ARTISTAS_DISCO_COMPACTO
     WHERE ARTISTA = 'Joni Mitchell'
       AND INVENTARIO_DISCO_COMPACTO.NOMBRE_CD =
         ARTISTAS_DISCO_COMPACTO.TITULO );
```

Si se ejecuta esta instrucción, se obtendrán los siguientes resultados de consulta:

NOMBRE_CD	EN_EXISTENCIA
Blue	42
Court and Spark	22

La mejor forma de comprender cómo funciona esta instrucción es observar cómo se evalúan las filas individuales. Como se aprenderá en el capítulo 12, las subconsultas como ésta son llamadas *subconsultas correlacionadas* debido a que la subconsulta se ejecuta para cada fila arrojada en la instrucción SELECT principal. Debido a que la cláusula WHERE de la subconsulta hace coincidir el valor NOMBRE_CD con el valor TITULO, el valor TITULO en la fila que se está eva-

luando (en la subconsulta) debe coincidir con el valor NOMBRE_CD para que esa fila pueda ser arrojada. Por ejemplo, la primera fila en la tabla INVENTARIO_DISCO_COMPACTO contiene un valor NOMBRE_CD de Famous Blue Raincoat. Cuando esta fila es comparada con el predicado EXISTS, el valor Famous Blue Raincoat se hace coincidir con el valor Famous Blue Raincoat de la columna TITULO en la tabla ARTISTAS_DISCO_COMPACTO. Además, el valor Joni Mitchell se hace coincidir con el valor ARTISTA para la fila Famous Blue Raincoat. Debido a que el valor ARTISTA es Jennifer Warnes, y no Joni Mitchell, la condición de búsqueda especificada en la cláusula WHERE de la subconsulta se evalúa como falsa; por lo tanto, no se arroja ninguna fila de subconsulta para la fila Famous Blue Raincoat. Como resultado, la cláusula WHERE en la instrucción SELECT principal se evalúa como falsa para la fila Famous Blue Raincoat de la tabla INVENTARIO_DISCO_COMPACTO, y la fila no se incluye en los resultados de la consulta.

Este proceso se repite para cada fila en la tabla INVENTARIO_DISCO_COMPACTO. Si la cláusula WHERE en la subconsulta se evalúa como verdadera, entonces el predicado EXISTS se evalúa como verdadero, lo que significa que la cláusula WHERE en la instrucción SELECT principal se evalúa como verdadera. En el caso de nuestro último ejemplo de la instrucción SELECT, solamente dos filas coinciden con este criterio.

NOTA

No tiene importancia cuáles columnas o cuántas columnas se especifican en la cláusula SELECT de la subconsulta en un predicado EXISTS. Este tipo de predicado se concentra solamente en si las filas serán arrojadas, en lugar de en el contenido de esas filas. Se puede especificar cualquier nombre de columna o simplemente un asterisco.

El predicado EXISTS, como se puede esperar, permite utilizar el inverso de la condición del predicado utilizando la palabra clave NOT:

```
SELECT *
  FROM INVENTARIO_DISCO_COMPACTO
 WHERE NOT EXISTS
       ( SELECT TITULO
         FROM ARTISTAS_DISCO_COMPACTO
        WHERE ARTISTA = 'Joni Mitchell'
          AND INVENTARIO_DISCO_COMPACTO.NOMBRE_CD =
            ARTISTAS_DISCO_COMPACTO.TITULO );
```

En este caso, todos los CD, *excepto* los CD de Joni Mitchell, se incluyen en los resultados de la consulta. Eso significa que si la cláusula WHERE de la subconsulta se evalúa como verdadera (lo que significa que la subconsulta arroja una fila), el predicado en sí mismo se evalúa como falso y no se arroja ninguna fila. Por otro lado, si la subconsulta no arroja una fila, el predicado se evalúa como verdadero y la fila se arroja en los resultados de la consulta de la instrucción SELECT principal.

Pregunta al experto

P: Se han proporcionado ejemplos que muestran que existe a menudo más de una forma para lograr el mismo resultado. ¿Cómo puede uno saber cuál opción seleccionar cuando se está escribiendo una instrucción SQL?

R: Encontrará que, mientras más aprende acerca de programación SQL y logra un mejor entendimiento de las facetas de cada instrucción, encontrará a menudo más de una forma de lograr los mismos resultados. En estos casos, su elección de los métodos dependerá a menudo de cuál instrucción sea la más sencilla de escribir o cuál realiza mejor el trabajo en una implementación SQL en particular. Al mismo tiempo que crece su comprensión de SQL, también lo hará su habilidad para elegir el método que sea mejor para cada situación. En muchos casos, la diferencia entre un método y otro no será muy grande, y su elección dependerá simplemente de sus preferencias personales. Sin embargo, también puede toparse con situaciones en las cuales la implementación de SQL en la que se está trabajando no soporte todos los métodos proporcionados en el estándar SQL. Por lo tanto, se debe seleccionar el método que pueda ser implementado en ese ambiente en particular. Cualquiera que sea el método que se utilice en cualquier ambiente dado, lo mejor por ahora es tener una base tan completa como sea posible acerca de los conceptos básicos de SQL. De esa forma se estará mejor preparado para las diferentes situaciones y mejor equipado para moverse de una implementación a otra. Adicionalmente, deberá aprender acerca de los temas de rendimiento relacionados a la implementación con la que se está trabajando. Deberá considerar los temas de rendimiento cuando se tome una decisión acerca de cuáles instrucciones SQL se deben utilizar.

P: Cuando se proporcionaron ejemplos del predicado EXISTS, las subconsultas siempre hacían coincidir columnas dentro de la cláusula WHERE de la subconsulta. ¿Esto es necesario?

R: Es posible, si se desea, crear un predicado EXISTS que no haga coincidir columnas en la subconsulta, como la siguiente instrucción:

```
SELECT TITULO, ARTISTA
FROM ARTISTAS_DISCO_COMPACTO
WHERE EXISTS
    ( SELECT NOMBRE_CD
      FROM INVENTARIO_DISCO_COMPACTO
      WHERE EN_EXISTENCIA > 10 );
```

En este caso, la subconsulta simplemente revisa si es que existe en la tabla INVENTARIO_DISCO_COMPACTO algún valor EN_EXISTENCIA mayor a 10. Si esa fila o filas existen, el predicado se evaluó como verdadero, lo que significa que la cláusula WHERE en la instrucción SELECT principal se evalúa como verdadera. Como resultado, se arrojan todas las filas en la tabla ARTISTAS_DISCO_COMPACTO. Utilizar este tipo de subconsulta generalmente no es muy útil debido a que ofrece pocas ventajas sobre una instrucción SELECT simple. Cuando se utiliza EXISTS, hacer coincidir columnas de diferentes tablas dentro de la subconsulta es esencial para proporcionar un filtrado significativo para la instrucción SELECT principal.

Determinar la cantidad de predicados de comparación

SQL incluye otro tipo de predicado denominado *predicado de comparación cuantificado*, que es un tipo de predicado utilizado en conjunción con un operador de comparación para determinar si *alguno* o *todos* los valores arrojados coinciden con los requerimientos de la búsqueda. SQL soporta tres predicados de comparación cuantificados: SOME, ANY y ALL. Se hace referencia a los predicados SOME y ANY como *cuantificadores existenciales*, y se encargan de comprobar si cualquier valor arrojado corresponde a los requisitos de la búsqueda. Estos dos predicados son idénticos en significado y pueden ser utilizados intercambiabilmente. El predicado ALL es llamado *cuantificador universal* y se ocupa de comprobar si *todos* los valores arrojados corresponden a los requisitos de la búsqueda. Ahora demos un vistazo más cercano a cada uno.

Utilizar los predicados SOME y ANY

Como se mencionó, los predicados SOME y ANY arrojan resultados idénticos. Para cada fila, los predicados comparan el valor en una columna especificada con los resultados de una subconsulta. Si la comparación se evalúa como verdadera para *cualquiera* de los resultados, la condición se toma como satisfactoria y se arroja esa fila. Para crear uno de estos predicados, se debe especificar el nombre de la columna que contiene los valores que se quieren comparar, el operador de comparación (véase la sección “Comparar datos SQL”), la palabra clave SOME o ANY, y la subconsulta. A pesar de que se pueden utilizar cualquiera de las palabras, particularmente yo prefiero utilizar ANY porque me parece más intuitiva, pero siéntase libre de utilizar la que guste.

Ahora veamos un ejemplo para ilustrar mejor cómo funcionan estos predicados. El ejemplo está basado en la tabla MENUDEO_CD y en la tabla REBAJA_CD, que se muestran en la figura 9-5.

En este ejemplo se busca consultar datos de la tabla REBAJA_CD. Se quiere arrojar sólo aquellas filas que tengan un valor MENUDEO menor que algunos de los valores MENUDEO en la tabla MENUDEO_CD.

MENUDEO_CD			REBAJA_CD	
NOMBRE_CD: VARCHAR(60)	MENUDEO: NUMERIC(5,2)	EN_EXISTENCIA: INT	TITULO: VARCHAR(60)	VENTA: NUMERIC(5,2)
Famous Blue Raincoat	16.99	5	Famous Blue Raincoat	14.99
Blue	14.99	10	Blue	12.99
Court and Spark	14.99	12	Court and Spark	14.99
Past Light	15.99	11	Past Light	14.99
Kojiki	15.99	4	Kojiki	13.99
That Christmas Feeling	10.99	8	That Christmas Feeling	10.99
Patsy Cline: 12 Greatest Hits	16.99	14	Patsy Cline: 12 Greatest Hits	16.99

Figura 9-5 Utilizar predicados de comparación cuantificados en las tablas MENUDEO_CD y REBAJA_CD.

Los valores RETAIL deberán ser de filas que tengan un valor EN_EXISTENCIA mayor a 9. En otras palabras, la consulta deberá arrojar solamente aquellos CD cuyo precio rebajado (REBAJA) sea menor que *cualquier* precio de lista (MENUDEO) en aquellos CD que haya una existencia mayor a nueve. Para cumplir este requisito se utilizará la siguiente instrucción SELECT:

```
SELECT TITULO, REBAJA
FROM REBAJA_CD
WHERE REBAJA < ANY
  ( SELECT MENUDEO
    FROM MENUDEO_CD
    WHERE EN_EXISTENCIA > 9 );
```

Si se prefiere, se puede utilizar la palabra clave SOME en lugar de la palabra clave ANY. Los resultados de la consulta serían los mismos, como se muestra en los siguientes:

TITULO	REBAJA
-----	-----
Famous Blue Raincoat	14.99
Blue	12.99
Court and Spark	14.99
Past Light	14.99
Kojiki	13.99
That Christmas Feeling	10.99

Ahora analicemos la instrucción SELECT más de cerca. El predicado ANY contiene la siguiente subconsulta:

```
SELECT RETAIL
FROM CD_RETAIL
WHERE EN_EXISTENCIA > 9
```

Si se ejecutara tal subconsulta por sí sola, se recibirían los siguientes resultados:

```
MENUDEO
-----
14.99
14.99
15.99
16.99
```

El valor REBAJA en cada fila en la tabla REBAJA_CD es por lo tanto comparado a los resultados de la subconsulta. Por ejemplo, la fila Past Light tiene un valor REBAJA de 14.99. Este valor es comparado a los resultados de la subconsulta para verificar si 14.99 es menor que *cualquier* valor. Debido a que éste es menor que 15.99 y 16.99, el predicado se evalúa como verdadero y la fila es arrojada. La única fila que no se evalúa como verdadera es la fila Patsy Cline: 12 Greatest Hits debido a que el valor REBAJA es 16.99, y éste no es menor a ninguno de los valores arrojados por los resultados de la consulta.

Puede utilizar cualquiera de los seis operadores de comparación en un predicado ANY o SOME. Por ejemplo, si se hubiera utilizado el operador Mayor que, solamente la fila Patsy Cline:

12 Greatest Hits habría sido arrojado debido a que hubiera sido la única fila con un valor REBAJA mayor que cualquier fila en los resultados de la subconsulta.

NOTA

Los predicados de comparación cuantificados no soportan una condición inversa como lo hacen otros predicados. En otras palabras, no se puede agregar la palabra NOT antes de ANY o SOME. Sin embargo, se pueden lograr los mismos resultados utilizando el operador desigual a (<>).

Utilizar el predicado ALL

El predicado ALL funciona muy parecido a los predicados SOME y ANY porque también compara valores de columna con los resultados de la subconsulta. Sin embargo, en lugar de que los valores de columna tengan que evaluarse como verdaderos para *cualquiera* de los valores resultantes, los valores de columna deben evaluarse como verdaderos para *todos* los valores resultantes; si no es así, la fila no será arrojada.

Regresemos al ejemplo anterior, solamente que esta vez se sustituirá la palabra clave ANY por la palabra clave ALL. La nueva instrucción SELECT lucirá de la siguiente manera:

```
SELECT TITULO, REBAJA
FROM REBAJA_CD
WHERE REBAJA < ALL
      ( SELECT MENUDEO
        FROM MENUDEO_CD
        WHERE EN_EXISTENCIA > 9 );
```

Si se ejecuta esta instrucción, se encontrará que los resultados de la consulta son bastante diferentes a lo que eran en el ejemplo anterior:

TITULO	REBAJA
Blue	12.99
Kojiki	13.99
That Christmas Feeling	10.99

Esta vez solamente se arrojaron tres filas debido a que son las únicas que cumplen la condición del predicado WHERE.

Si se da un vistazo más cercano a la instrucción, se encontrará que la subconsulta arroja los mismos valores que en los ejemplos anteriores. Sin embargo, el valor SALE para cada fila en la tabla REBAJA_CD debe ahora ser menor que todos los valores en los resultados de la subconsulta. Por ejemplo, la fila Kojiki contiene un valor REBAJA de 13.99. Los resultados de la subconsulta incluyen los valores 14.99, 15.99 y 16.99. El valor 13.99 es menor que todos los tres valores resultantes de la subconsulta, lo cual significa que el predicado se evalúa como verdadero y que esa fila se incluye en los resultados de la consulta. Por otro lado, la fila Past Light contiene un valor REBAJA de 14.99, el cual no es menor que el valor de la subconsulta 14.99, por lo que esa fila no se incluye en los resultados de la consulta.

Pregunta al experto

P: En los análisis acerca de los predicados de comparación cuantificados, se incluyeron ejemplos sobre cómo utilizar estos predicados; sin embargo, los ejemplos incluían solamente un predicado en la cláusula WHERE. ¿Es posible utilizar múltiples predicados cuando se utiliza un predicado de comparación cuantificado?

R: Sí. Es posible utilizar múltiples predicados. Al igual que con cualquier otro tipo de predicado, simplemente deben conectarse los predicados utilizando la palabra clave AND o la palabra clave OR. Pero debe asegurarse de que la lógica que se está utilizando no solamente tenga sentido en términos de los datos que serán arrojados, sino también que tenga sentido para que pueda comprenderse la instrucción por sí misma. Como resultado, la mejor forma de tratar con este tipo de situaciones es destacando cada predicado entre paréntesis y luego conectando estas expresiones parentéticas con AND u OR. Por ejemplo, supongamos que se quiere utilizar el ejemplo en la sección “Utilizar los predicados SOME y ANY” y agregarle un predicado LIKE. (El ejemplo está basado en la figura 9-5.) Se puede crear una instrucción SELECT similar a la siguiente:

```
SELECT TITULO, SALE
FROM REBAJA_CD
WHERE ( REBAJA < ANY ( SELECT MENUDEO
                        FROM REBAJA_CD
                        WHERE EN_EXISTENCIA > 9 ) )
AND ( TITULO LIKE ('%Blue%') );
```

Observe que cada predicado ha sido encerrado en un conjunto de paréntesis y que éstos están unidos por AND. Si se ejecuta esta instrucción, los resultados de la consulta acatarán la condición del predicado ANY y del predicado LIKE, la cual especifica que el valor TITLE incluya la palabra Blue. Si se prefiere, se pueden escribir estas instrucciones sin encerrar los predicados entre paréntesis, pero entonces las instrucciones pueden empezar a confundirse, y en estructuras más complejas pueden también empezar a producir resultados inesperados.

Al igual que con los predicados ANY y SOME, es posible utilizar cualquiera de los seis operadores de comparación con el predicado ALL. Adicionalmente, se puede crear cualquier tipo de subconsulta, siempre y cuando encaje lógicamente con la instrucción SELECT principal. El punto que debe recordarse es que el valor de la columna debe ser verdadero para todos los resultados de la subconsulta, y no solamente para algunos de ellos.

Pruebe esto 9-2 Utilizar subconsultas en predicados

Este ejercicio básicamente comienza donde se quedó el anterior. Una vez más, trabajará con predicados, sólo que esta vez será con aquellos que utilizan subconsultas. Son estos los predicados que se analizaron desde el último ejercicio. Incluyen a los predicados IN, EXISTS, ANY y ALL. Al igual que con el ejercicio anterior, se aplicarán estos predicados a las tablas que se crearon en la base de datos INVENTARIO. Puede descargar el archivo Try_This_09.txt (en inglés), que contiene las instrucciones SQL utilizadas en este ejercicio.

Paso a paso

1. Abra la aplicación de cliente para sus RDBMS y conéctese con la base de datos INVENTARIO.
2. En su primera instrucción utilizará un predicado IN para consultar datos de la tabla DISCOS_COMPACTOS. Se quiere ver la información de CD y de inventario para los CD publicados por la compañía Decca Record Company. Para averiguar cuáles son estos CD, se debe crear una subconsulta que consulte los datos de la tabla DISQUERAS_CD. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT TITULO_CD, EN_EXISTENCIA
FROM DISCOS_COMPACTOS
WHERE ID_DISQUERA IN
      ( SELECT ID_DISQUERA
        FROM DISQUERAS_CD
        WHERE NOMBRE_COMPAÑIA = 'Decca Record Company' );
```

Los resultados de la consulta deberán incluir solamente dos filas. Ambas filas tendrán un valor ID_DISQUERA de 833, que es el valor arrojado por la subconsulta.

3. Ahora se intentará una instrucción SELECT similar a la del paso 2, sólo que esta vez se utilizará un predicado EXISTS para arrojar los datos. Adicionalmente, tendrá que agregar un predicado a la cláusula WHERE de la subconsulta que hará coincidir el valor ID_DISQUERA en la tabla DISCOS_COMPACTOS con el valor ID_DISQUERA en la tabla DISQUERAS_CD. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT TITULO_CD, EN_EXISTENCIA FROM DISCOS_COMPACTOS
WHERE EXISTS
      ( SELECT ID_DISQUERA FROM DISQUERAS_CD
        WHERE DISCOS_COMPACTOS.ID_DISQUERA = DISQUERAS_CD.ID_DISQUERA
          AND ID_DISQUERA > 830 );
```

Observe que uno de los predicados en la cláusula WHERE de la subconsulta utiliza un operador de comparación para buscar los valores ID_DISQUERA mayores a 830. Si se busca en la tabla DISQUERAS_CD, se encontrará que hay seis filas que contienen valores ID_DISQUE-

RA mayores a 830. Entonces, si se fuera a hacer coincidir estos seis valores con los valores ID_DISQUERA en la tabla DISCOS_COMPACTOS, se encontraría que 11 filas se evaluarían como verdaderas. Éstas serán las 11 filas arrojadas por la instrucción SELECT.

4. En esta instrucción se utilizará un predicado ANY para comparar los valores ID_DISQUERA en la tabla DISQUERAS_CD con los valores ID_DISQUERA en la tabla DISCOS_COMPACTOS que están incluidos en filas con un valor EN_EXISTENCIA mayor a 20. Los valores ID_DISQUERA en la tabla DISQUERAS_CD pueden coincidir con cualquier valor en los resultados de la subconsulta. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT ID_DISQUERA, NOMBRE_COMPANIA
FROM DISQUERAS_CD
WHERE ID_DISQUERA = ANY
      ( SELECT ID_DISQUERA
        FROM DISCOS_COMPACTOS
        WHERE EN_EXISTENCIA > 20 );
```

La consulta deberá arrojar solamente cinco filas.

5. Ahora trate de crear la misma instrucción SELECT del paso 4, sólo que esta vez utilizando un predicado ALL en lugar de un predicado ANY. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT ID_DISQUERA, NOMBRE_COMPANIA
FROM DISQUERAS_CD
WHERE ID_DISQUERA = ALL
      ( SELECT ID_DISQUERA
        FROM DISCOS_COMPACTOS
        WHERE EN_EXISTENCIA > 20 );
```

Encontrará que con esta consulta no se arroja ninguna fila. Esto se debe a que la subconsulta arroja ocho filas con cinco diferentes valores. El valor ID_DISQUERA para cada fila en la tabla DISQUERAS_CD no puede hacer coincidir todos los valores, solamente uno o algunos de ellos. La única forma en que se arrojaría alguna fila en este caso sería si la subconsulta arroja solamente una fila o múltiples filas, todas con el mismo valor.

6. Ahora intente modificar la instrucción SELECT cambiando el predicado de comparación en la cláusula WHERE de la subconsulta a mayor que 40. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT ID_DISQUERA, NOMBRE_COMPANIA
FROM DISQUERAS_CD
WHERE ID_DISQUERA = ALL
      ( SELECT ID_DISQUERA
        FROM DISCOS_COMPACTOS
        WHERE EN_EXISTENCIA > 40 );
```

Los resultados de la consulta ahora arrojarán una fila. Esto se debe a que la subconsulta arroja sólo una fila que cumple la condición del predicado ALL.

7. Cierre la aplicación de cliente.

Resumen de Pruebe esto

En este ejercicio se utilizaron los predicados IN, EXISTS, ANY y ALL para consultar datos de la base de datos INVENTARIO. También pudo haberse utilizado el predicado SOME en lugar del predicado ANY. Combinados con los pasos del ejercicio anterior, las instrucciones aquí deberán haberle permitido probar con una gran variedad de predicados. Según vaya aprendiendo más acerca de las subconsultas, será capaz de crear predicados incluso más elaborados, unos que pueda utilizar no solamente en las instrucciones SELECT, sino también en las instrucciones UPDATE y DELETE. Entretanto, se le sugiere experimentar con diferentes tipos de instrucciones SELECT e intentar diferentes predicados dentro de esas instrucciones para ver exactamente qué tipos de resultados de consulta pueden arrojar.

Autoexamen Capítulo 9

1. ¿En cuál cláusula de la instrucción SELECT se incluyen predicados?
2. ¿Cuál símbolo de operador de comparación deberá utilizarse para expresar una condición desigual?
 - A <=
 - B >=
 - C <>
 - D =<
3. ¿Cuáles palabras clave pueden utilizarse para combinar predicados en una cláusula WHERE?
4. Se quiere consultar una tabla que incluye la columna PRECIO. Es necesario asegurarse que todas las filas arrojadas tengan un valor PRECIO de 13.99. ¿Cuál predicado deberá utilizarse?
5. Se crea la siguiente instrucción SQL:

```
SELECT TITULO_CD, PRECIO_MENUDEO
FROM CDS_A_LA_MANO
WHERE PRECIO_MENUDEO >= 14
AND PRECIO_MENUDEO <= 16;
```

¿Qué predicado puede utilizarse en lugar de los dos predicados mostrados en esta instrucción?
6. ¿Qué palabra clave puede agregarse a un predicado BETWEEN para encontrar el inverso de la condición especificada por el predicado?
7. ¿Cuándo se utiliza un valor nulo en una columna?
8. Se quiere consultar una tabla para determinar cuáles valores son nulos. ¿Qué tipo de predicado deberá utilizarse?

9. Se está creando una instrucción SELECT que consulta la tabla BIO_ARTISTAS. Se quieren arrojar todas las columnas en la tabla, pero se quieren arrojar sólo aquellas columnas que no contengan valores nulos en la columna LUGAR_DE_NACIMIENTO. ¿Cuál instrucción SELECT deberá utilizarse?
10. Se está consultando la tabla INVENTARIO_CD. Se quiere ver todas las columnas, pero se requiere ver sólo aquellas columnas que contengan la palabra Christmas en el nombre del CD. Los nombres están almacenados en la columna TITULO_CD. ¿Cuál instrucción SELECT deberá utilizarse?
11. ¿Cuál es la diferencia entre un signo de porcentaje y un guión bajo cuando se usan en un predicado LIKE?
12. ¿Cuáles dos tipos de fuentes de datos pueden utilizarse en un predicado IN?
13. ¿Cuál tipo de predicado se ocupa solamente de determinar si una subconsulta arroja cualquier fila o no?
14. ¿Cuáles nombres de columna deben ser especificados en un predicado EXISTS?
15. Se está creando una instrucción SELECT que incluye un predicado en la cláusula WHERE. Se quiere utilizar un operador de comparación para comparar los valores en una de las columnas con los resultados de una subconsulta. Se quiere que el predicado se evalúe como verdadero para cualquiera de los resultados de la subconsulta. ¿Qué tipo de predicado deberá utilizarse?
 - A EXISTS
 - B ANY
 - C ALL
 - D IN
16. ¿Cuál es la diferencia entre un predicado SOME y un predicado ANY?
17. ¿Cómo difiere el predicado ALL del predicado SOME?

Capítulo 10

Trabajar con funciones
y expresiones de valor

Habilidades y conceptos clave

- Utilizar funciones Set
 - Utilizar funciones de valor
 - Utilizar expresiones de valor
 - Utilizar valores especiales
-

En partes anteriores de este libro se han presentado brevemente varias funciones y expresiones relacionadas con valores. Estos valores y expresiones se utilizan en ejemplos y ejercicios en varios capítulos para demostrar diferentes componentes de SQL. En este capítulo se dará un vistazo más cercano a muchos de estos valores y expresiones, centrándose en aquellos que más probablemente pueda utilizar un programador principiante de SQL. Deberá conservar en mente, sin embargo, que este capítulo cubre solamente una porción de los muchos tipos de funciones y expresiones soportados por SQL. Adicionalmente, las implementaciones de SQL pueden variar respecto a cuáles funciones y expresiones SQL pueden soportar, cómo son implementados esos valores y expresiones, y cuáles funciones y expresiones estándar y no estándar incluyen en sus productos. Asegúrese de revisar la documentación del producto para determinar cuáles funcionalidades son soportadas. En general, se incluyen en este capítulo aquellas funciones y expresiones más comúnmente soportadas por las implementaciones SQL.

Utilizar funciones Set

En el capítulo 3 se introdujo el concepto de función. Como se puede recordar, una función es una operación nombrada que realiza tareas definidas que normalmente no se pueden realizar utilizando solamente instrucciones SQL. Es un tipo de rutina que toma parámetros de entrada, los cuales están encerrados en paréntesis, y arroja valores basados en esos paréntesis. Una propiedad importante de las funciones es que cada ejecución de una función arroja exactamente un valor de datos, y ésta es la razón por la que las funciones pueden ser utilizadas en lugar de los nombres de columna de una tabla en la lista SELECT de una consulta (la función arroja un valor único para cada fila procesada por la consulta). Ya hemos visto ejemplos de funciones como SUM y AVG. Ambas funciones son conocidas como funciones set. Una *función set*, a veces nombrada función *agregada*, procesa o calcula datos y arroja los valores apropiados. Las funciones set requieren que los datos estén agrupados de cierta manera, como sería el caso si la cláusula GROUP BY fuera utilizada en una instrucción SELECT. Si las filas no están agrupadas explícitamente de esa manera, la tabla completa es tratada como un grupo.

En esta sección se analizan cinco funciones set: COUNT, MAX, MIN, SUM y AVG. Todas estas funciones son comúnmente soportadas en las implementaciones SQL. Para todas las funciones set, se proporcionan ejemplos acerca de cómo se utilizarían en la cláusula SELECT de una instrucción SELECT. Los ejemplos están basados en la tabla mostrada en la figura 10-1.

NOMBRE_ARTISTA: VARCHAR(60)	NOMBRE_CD: VARCHAR(60)	VENDIDOS: INT
Jennifer Warnes	Famous Blue Raincoat	23
Joni Mitchell	Blue	45
Joni Mitchell	Court and Spark	34
William Ackerman	Past Light	12
Bing Crosby	That Christmas Feeling	34
Patsy Cline	Patsy Cline: 12 Greatest Hits	54
John Barry	Out of Africa	23
Leonard Cohen	Leonard Cohen The Best of	20
Bonnie Raitt	Fundamental	29
B.B. King	Blues on the Bayou	18

Figura 10-1 Utilizar las funciones set en la tabla CDS_ARTISTA.

Utilizar la función COUNT

La primera función set que veremos será la función COUNT. Como su nombre sugiere, la función COUNT cuenta el número de filas en una tabla o el número de valores en una columna, según se especifique en la instrucción SELECT. Cuando se utiliza la función COUNT, se debe especificar un nombre de columna para contar el número de valores que no sean nulos en una columna, o un asterisco para contar todas las filas en una tabla independientemente de los valores nulos. Por ejemplo, si se quiere saber el número total de filas en la tabla CDS_ARTISTA, se puede utilizar la siguiente instrucción SELECT:

```
SELECT COUNT(*) AS FILAS_TOTALES
FROM CDS_ARTISTA;
```

En esta instrucción, la función COUNT es utilizada con un asterisco (en paréntesis) para contar cada fila en la tabla CDS_ARTISTA y arrojar la cuenta total. El valor arrojado se enlista en la columna FILAS_TOTALES, que es el nombre dado a la columna arrojada en los resultados de la consulta, como se muestra en los siguientes resultados:

```
FILAS_TOTALES
-----
10
```

Como se puede ver, los resultados de la consulta incluyen solamente un valor (una fila con una columna), como podría esperarse de una función set utilizada sin ningún agrupamiento de fila. El valor 10 indica que la tabla CDS_ARTISTA contiene 10 filas.

Al igual que con cualquier otra clase de instrucción `SELECT`, se pueden cualificar los resultados de la consulta agregando a la instrucción las cláusulas necesarias. Por ejemplo, supongamos que se quiere averiguar cuántas filas incluye un valor `VENDIDOS` mayor a 20. Se puede modificar la instrucción `SELECT` para incluir una cláusula `WHERE`:

```
SELECT COUNT(*) AS FILAS_TOTALES
FROM CDS_ARTISTA
WHERE VENDIDOS > 20;
```

El valor arrojado ahora será 7 en lugar de 10, debido a que solamente siete filas cumplen con la condición de búsqueda especificada en la cláusula `WHERE`.

Se puede encontrar con que, en lugar de consultar el número de filas en una tabla, se quiera conocer el número de valores en una columna dada (sin incluir los valores nulos). En este caso, se especificaría el nombre de la columna en lugar del asterisco (*). Por ejemplo, supongamos que se modifica la instrucción `SELECT` mostrada en el último ejemplo para contar valores en la columna `NOMBRE_ARTISTA`:

```
SELECT COUNT(NOMBRE_ARTISTA) AS TOTAL_DE_ARTISTAS
FROM CDS_ARTISTA
WHERE VENDIDOS > 20;
```

Cuando se ejecuta esta consulta, el valor arrojado una vez más es 7. Esto significa que siete valores `NOMBRE_ARTISTA` tienen un valor `VENDIDOS` mayor a 20. Sin embargo, esta instrucción no contabiliza aquellos valores `NOMBRE_ARTISTA` que pudieran estar duplicados. Si se quiere llegar a un conteo que tome en consideración los valores duplicados, se puede agregar la palabra clave `DISTINCT` a la función `COUNT`:

```
SELECT COUNT(DISTINCT NOMBRE_ARTISTA) AS TOTAL_DE_ARTISTAS
FROM CDS_ARTISTA
WHERE VENDIDOS > 20;
```

Esta vez el valor arrojado es de 6 en lugar de 7. Esto se debe a que la columna `NOMBRE_ARTISTA` incluye dos instancias del valor Joni Mitchell. La columna contiene solamente seis valores únicos que cumplen la condición establecida en el criterio de búsqueda.

NOTA

Tenga en cuenta que la instrucción `SELECT` es procesada en un orden específico: primero la cláusula `FROM`, luego la cláusula `WHERE`, y después la cláusula `SELECT`. Como resultado, la función `COUNT` aplica solamente a las filas que cumplen la condición de búsqueda definida en la cláusula `WHERE`. Las filas que no estén incluidas en los resultados de la cláusula `WHERE` no tendrán relevancia en la función `COUNT`. Para mayor información acerca de la instrucción `SELECT`, véase el capítulo 7.

Como ya se mencionó, si la columna especificada en la función aún contiene valores nulos, esos valores no estarán incluidos en la cuenta. Por ejemplo, si se agregara una fila a la tabla `CDS_ARTISTA` con un valor `NOMBRE_ARTISTA` nulo y un valor `VENDIDOS` mayor a 20, la instrucción `SELECT` mostrada en el ejemplo anterior aún arrojaría un valor de 6 debido a que el valor nulo no sería contado. Sin embargo, si se utilizara un asterisco en lugar de un nombre de columna en la función `COUNT`, todas las filas serían contadas, incluso si algunas tuvieran valores nulos.

Utilizar las funciones MAX y MIN

Las funciones MAX y MIN son tan similares que resulta ideal revisarlas juntas. La función MAX arroja el valor más alto para la columna especificada, y la función MIN arroja el valor más bajo. Ambas funciones requieren que se especifique un nombre de columna. Por ejemplo, supongamos que se quiere arrojar el valor más alto de la columna VENDIDOS en la tabla CDS_ARTISTA. La instrucción SELECT quedaría de la manera siguiente:

```
SELECT MAX(VENDIDOS) AS MAX_VENDIDOS
FROM CDS_ARTISTA;
```

Cuando se ejecuta esta instrucción, los resultados de la consulta incluirán solamente un valor (una fila y una columna), como se muestra en los siguientes resultados:

```
MAX_VENDIDOS
-----
54
```

Este resultado, por sí mismo, no resulta de mucha ayuda. Sería mejor si los resultados de la consulta también incluyeran el nombre del artista y el CD. Sin embargo, SQL no soporta una instrucción SELECT como la siguiente:

```
SELECT NOMBRE_ARTISTA, NOMBRE_CD, MAX(VENDIDOS)
FROM CDS_ARTISTA;
```

Debido a que las funciones set tratan los datos como grupos, no es posible especificar el nombre del artista y el nombre del CD sin agrupar los datos de alguna forma. De la manera en que está ahora, la función MAX trata la tabla entera como un grupo; sin embargo, ni los valores de NOMBRE_ARTISTA ni los de NOMBRE_CD están agrupados de ninguna forma, por lo que la cláusula SELECT se vuelve ilógica.

Cada vez que se incluya una función set en una instrucción SQL, cada argumento en la lista SELECT deberá ser una función set o estar incluido en un grupo (utilizando la cláusula GROUP BY descrita posteriormente en este tema). Una forma relativa a esto es utilizar una subconsulta en la cláusula WHERE para arrojar el valor máximo y luego arrojar la información necesaria pasada en ese valor, como se muestra en el siguiente ejemplo:

```
SELECT NOMBRE_ARTISTA, NOMBRE_CD, VENDIDOS
FROM CDS_ARTISTA
WHERE VENDIDOS = ( SELECT MAX(VENDIDOS)
                   FROM CDS_ARTISTA );
```

La subconsulta encuentra el valor máximo (54) y utiliza ese valor como una condición en la cláusula WHERE. El valor VENDIDOS debe ser igual a 54 mientras ése sea el valor más alto de VENDIDOS en la tabla. Una vez que se define la condición de búsqueda necesaria en la cláusula WHERE, se pueden utilizar esos resultados para arrojar la información que se necesita. Si se ejecuta esta instrucción, se arrojará solamente una fila:

NOMBRE_ARTISTA	NOMBRE_CD	VENDIDOS
-----	-----	-----
Patsy Cline:	Patsy Cline 12 Greatest Hits	54

Como se puede ver, ahora se tiene toda la información necesaria para determinar cuál artista y cuáles CD han tenido las ventas más altas.

Como se dijo anteriormente, las funciones MAX y MIN son muy similares. Si se reemplaza MAX por MIN en el ejemplo anterior, los resultados de la consulta serán los siguientes:

NOMBRE_ARTISTA	NOMBRE_CD	VENDIDOS
-----	-----	-----
William Ackerman	Past Light	12

Se arroja la fila Past Light debido a que es la fila con el valor VENDIDOS más bajo.

Las funciones MAX y MIN no están limitadas a datos numéricos. También pueden ser utilizadas para comparar cadenas de caracteres. Por ejemplo, supongamos que se quiere saber cuál artista aparecería primero alfabéticamente. La siguiente instrucción arrojará B.B. King:

```
SELECT MIN(NOMBRE_ARTISTA) AS LOW_NOMBRE
FROM CDS_ARTISTA;
```

Si se utiliza la función MAX, la instrucción arrojará William Ackerman.

NOTA

Es muy probable que las tablas en su base de datos agrupen separadamente los nombres de los apellidos debido a que ése es un diseño más flexible. Aquí se han incluido ambos nombres a una sola columna para proporcionarle ejemplos simples sobre cómo funcionan las diferentes instrucciones. Si los nombres estuvieran separados en dos columnas, la función MIN o MAX necesitaría ser utilizada en la columna apropiada.

Ahora volvamos un poco a la idea de la agrupación de datos. Como se mencionó, una función set trata a una tabla como un grupo si ningún agrupamiento ha sido implementado. Sin embargo, se puede utilizar fácilmente una cláusula GROUP BY para agrupar los datos. Supongamos que se quiere conocer la cantidad máxima vendida por cada artista. Se pueden agrupar los datos basados en los valores NOMBRE_ARTISTA:

```
SELECT NOMBRE_ARTISTA, MAX(VENDIDOS) AS MAX_VENDIDOS
FROM CDS_ARTISTA
WHERE VENDIDOS > 30
GROUP BY NOMBRE_ARTISTA;
```

La cláusula WHERE arroja solamente aquellas filas con un valor VENDIDOS mayor a 30. Luego las filas son agrupadas de acuerdo con los valores NOMBRE_ARTISTA. Una vez que han sido agrupadas, la cantidad máxima se arroja para cada artista, como se muestra en los siguientes resultados de la consulta:

NOMBRE_ARTISTA	MAX_VENDIDOS
-----	-----
Bing Crosby	34
Joni Mitchell	45
Patsy Cline	54

La cláusula GROUP BY crea tres grupos, uno para cada artista que cumple con la condición de la búsqueda definida en la cláusula WHERE. De estos tres grupos, solamente uno consta de valores duplicados: Joni Mitchell. Debido a que hay dos filas Joni Mitchell en la tabla CDS_ARTISTA, también hay dos valores VENDIDOS: 45 y 34. Como se puede ver, el valor más alto es 45, que es el que está incluido en los resultados de la consulta para el grupo Joni Mitchell. Si la función MIN hubiera sido utilizada en la instrucción SELECT, sería el valor 34 el que hubiera sido arrojado. Para los otros dos grupos de artistas, debido a que sólo existe un valor para cada uno, es ese mismo valor el que es arrojado sin importar si se utiliza la función MAX o la función MIN.

Utilizar la función SUM

A diferencia de las funciones MIN y MAX, que seleccionan los valores más alto y más bajo de una columna, la función SUM agrupa valores de columna. Esto es particularmente útil cuando se necesita encontrar los totales para datos agrupados (a pesar de que la función SUM, al igual que cualquier otra función set, trata a la tabla entera como un grupo único si ningún dato ha sido explícitamente agrupado).

Para entender mejor la función SUM, tomemos el último ejemplo que vimos y hagámosle algunas modificaciones:

```
SELECT NOMBRE_ARTISTA, SUM(VENDIDOS) AS TOTAL_VENDIDOS
FROM CDS_ARTISTA
WHERE VENDIDOS > 30
GROUP BY NOMBRE_ARTISTA;
```

Como se vio anteriormente, la cláusula WHERE arroja solamente aquellas filas con un valor VENDIDOS mayor a 30. Luego estas filas son agrupadas de acuerdo con los valores de NOMBRE_ARTISTA. Una vez agrupadas, las cantidades totales de cada grupo ARTIST son arrojadas en los resultados de la consulta:

NOMBRE_ARTISTA	TOTAL_VENDIDOS
Bing Crosby	34
Joni Mitchell	79
Patsy Cline	54

Observe que los resultados de la consulta incluyen los mismos tres grupos que fueron arrojados en el ejemplo anterior. La única diferencia es que el valor TOTAL_VENDIDOS en la fila Joni Mitchell es de 79, a diferencia de 45 o 34. La función SUM junta estos dos valores y arroja un valor de 79. Debido a que los otros dos grupos solamente constan de una entrada, sus valores TOTAL_VENDIDOS son los mismos que sus valores VENDIDOS en la tabla CDS_ARTISTA.

No es obligatorio utilizar una cláusula GROUP BY en una instrucción SELECT que utilice una función SUM. Se puede crear una instrucción SELECT tan simple como la siguiente:

```
SELECT SUM(VENDIDOS) AS TOTAL_VENDIDOS
FROM CDS_ARTISTA;
```

Esta instrucción simplemente coloca juntos todos los valores en la columna VENDIDOS y arroja un valor de 292. Por sí misma, esta información no siempre es de mucha ayuda, y es por eso que utilizar la función junto con una cláusula GROUP BY resulta mucho más efectivo.

Utilizar la función AVG

Como se puede imaginar, la función AVG simplemente promedia los valores en una columna especificada. Al igual que la función SUM, es más efectiva cuando se utiliza junto con una cláusula GROUP BY, a pesar de que puede ser utilizada sin la cláusula, como se muestra en el siguiente ejemplo:

```
SELECT AVG(VENDIDOS) AS PROM_VENDIDOS
FROM CDS_ARTISTA;
```

Esta instrucción arroja un valor de 29, que está basado en los valores VENDIDOS en la tabla CDS_ARTISTA. Esto significa que, para todos los CD enlistados en la tabla, se ha vendido un promedio de 29 de cada uno. A pesar de que esta información puede ser bastante útil, sería de mucha más ayuda si se estableciera una instrucción que agrupe los datos:

```
SELECT NOMBRE_ARTISTA, AVG(VENDIDOS) AS PROM_VENDIDOS
FROM CDS_ARTISTA
WHERE VENDIDOS > 30
GROUP BY NOMBRE_ARTISTA;
```

Si se ejecuta esta instrucción, se obtendrán los siguientes resultados de la consulta:

NOMBRE_ARTISTA	PROM_VENDIDO
Bing Crosby	34
Joni Mitchell	39
Patsy Cline	54

Como en los ejemplos anteriores, se crearon tres grupos, y para cada grupo se calcula un promedio basado en los valores de la columna VENDIDOS. Para la fila Joni Mitchell, este promedio está basado en los valores VENDIDOS de 45 y 34. Para las otras dos filas, el promedio es el mismo que el valor VENDIDOS debido a que solamente hay una fila por cada artista.

NOTA

La precisión de los valores arrojados por la función AVG depende del tipo de datos de la columna, si se utilizan decimales, y de cómo promedia los números la implementación SQL. Por ejemplo, el promedio exacto para la fila Joni Mitchell es 39.5, pero debido a que la columna VENDIDOS está configurada con un tipo de datos INT, solamente se utilizan números enteros. Para algunas implementaciones, el .5 es ignorado y no se redondea, como se muestra en los últimos resultados de consulta de ejemplo.

Utilizar funciones de valor

Las funciones de valor son un tipo de función que permite arrojar un valor que de alguna manera calcula o deriva información de los datos almacenados dentro de las tablas o de la misma imple-

mentación SQL. Las funciones de valor son similares a las funciones set en el sentido de que realizan algún tipo de acción por debajo del agua para llegar a ese valor. Sin embargo, las funciones de valor son diferentes de las funciones set en que no requieren que los datos sean agrupados.

SQL soporta diferentes funciones de valor. Cuáles funciones son soportadas por cuáles implementaciones SQL puede variar ampliamente. Además, el significado de un nombre de función puede a veces variar de una implementación a otra. Aun con eso, existen algunas consistencias entre las diferentes implementaciones, y esas son las funciones de valor en las cuales nos enfocaremos.

Las funciones de valor que se discutirán se dividen en dos categorías: funciones de valor de cadena y funciones de valor fecha y hora. Para ilustrar cómo trabajan estas funciones se utilizará la tabla FECHAS_VENTAS, mostrada la figura 10-2.

Trabajar con funciones de valor de cadena

Una función de valor de cadena permite manipular datos de cadenas de caracteres para producir un valor preciso que esté basado en la cadena de caracteres original. Cuando se utiliza una función de valor de cadena, se debe proporcionar la cadena de caracteres como un parámetro de la función. Ese parámetro es entonces convertido a un nuevo valor de acuerdo con el propósito de esa función y de otros parámetros que pudieran ser especificados. En esta sección se presentarán tres funciones de valor de cadena: SUBSTRING, UPPER y LOWER.

Utilizar la función de valor de cadena SUBSTRING

La función del valor de cadena SUBSTRING extrae un número definido de caracteres de una cadena de caracteres identificada para crear una nueva cadena. Esa cadena de caracteres original puede ser derivada de una columna o puede ser declarada explícitamente. En ambos casos, la cade-

DISCO_COMPACTO: VARCHAR(60)	FECHA_VENTA: TIMESTAMP
Famous Blue Raincoat	2002-12-22 10:58:05.120
Blue	2002-12-22 12:02:05.033
Court and Spark	2002-12-22 16:15:22.930
Past Light	2002-12-23 11:29:14.223
That Christmas Feeling	2002-12-23 13:32:45.547
Patsy Cline: 12 Greatest Hits	2002-12-23 15:51:15.730
Out of Africa	2002-12-23 17:01:32.270
Leonard Cohen The Best of	2002-12-24 10:46:35.123
Fundamental	2002-12-24 12:19:13.843
Blues on the Bayou	2002-12-24 14:15:09.673

Figura 10-2 Uso de funciones de valor en la tabla FECHAS_VENTAS.

na de caracteres pasa como un parámetro de la función SUBSTRING, junto con un punto de inicio y, opcionalmente, una especificación de longitud. Por ejemplo, supongamos que se quieren arrojar solamente los primeros 10 caracteres de los valores en la columna DISCO_COMPACTO en la tabla FECHAS_VENTAS. Se puede crear una instrucción SELECT similar a la siguiente:

```
SELECT SUBSTRING(DISCO_COMPACTO FROM 1 FOR 10) AS NOMBRE_ABREVIADO
FROM FECHAS_VENTAS;
```

La función SUBSTRING incluye tres parámetros. El primero es el nombre de la columna, DISCO_COMPACTO, que identifica la fuente utilizada para la cadena de caracteres. El siguiente parámetro, FROM 1, indica que la función empezará a contar a partir del primer carácter. El tercer parámetro, 10, sigue a la palabra clave FOR. El parámetro FOR 10, que es opcional, indica que serán incluidos hasta 10 caracteres en la nueva cadena de caracteres.

NOTA

La mayoría de las implementaciones, incluyendo SQL Server, MySQL y Oracle, no utilizan las palabras clave FROM y FOR (simplemente se separan los parámetros utilizando comas). Además, en Oracle la función es llamada SUBSTR. Aquí está la misma instrucción modificada para Oracle:

```
SELECT SUBSTR(DISCO_COMPACTO, 1, 10) AS NOMBRE_ABREVIADO
FROM FECHAS_VENTAS;
```

Si se ejecuta esta instrucción SELECT, se obtendrán los siguientes resultados de la consulta:

```
NAME_ABREVIADO
```

```
-----
```

```
Famous Blu
Blue
Court and
Past Light
That Chris
Patsy Clin
Out of Afr
Leonard Co
Fundamenta
Blues on t
```

Observe que solamente los primeros 10 caracteres de cada valor DISCO_COMPACTO están incluidos en los resultados. Para aquellos valores que tiene menos de 10 caracteres, aparece el nombre completo.

El parámetro FROM puede aceptar un número negativo o un cero como parámetro, asumiendo que la implementación SQL lo permite. Cuando utilice un número negativo o un cero, tenga en cuenta que el 1 representa lo que se puede considerar como una posición de inicio normal. El siguiente carácter a la izquierda del 1 es 0. El carácter a la izquierda de 0 es -1, y así sucesivamente. El parámetro FOR cuenta caracteres empezando desde el punto de inicio. Si se utiliza un cero o un número negativo, la función SUBSTRING actúa como si de todas maneras existieran caracteres

en esos lugares. Por ejemplo, supongamos que se modifica la instrucción SELECT anterior de la siguiente manera:

```
SELECT SUBSTRING(DISCO_COMPACTO FROM -2 FOR 10) AS NOMBRE_ABREVIADO
FROM FECHAS_VENTAS;
```

Si se ejecuta esta instrucción, solamente los primeros siete caracteres de cada nombre serían arrojados. Si en lugar de eso se utiliza un cero, solamente los primeros nueve caracteres serían arrojados. Solamente cuando se utiliza un parámetro FROM de 1 se arroja exactamente el número de caracteres (desde la cadena de caracteres) que están especificados por el parámetro FOR.

La función SUBSTRING no está limitada a la cláusula SELECT. De hecho, utilizarla en una cláusula WHERE puede ser muy útil cuando se define una condición de búsqueda. Por ejemplo, la siguiente instrucción SELECT utiliza la función SUBSTRING para arrojar filas que inician con Blue:

```
SELECT DISCO_COMPACTO, FECHA_VENTA
FROM FECHAS_VENTAS
WHERE SUBSTRING(DISCO_COMPACTO FROM 1 FOR 4) = 'Blue';
```

En esta instrucción, la función SUBSTRING arroja los primeros cuatro caracteres de los valores DISCO_COMPACTO y los compara con el valor Blue. Solamente dos filas son incluidas en los resultados de la consulta:

DISCO_COMPACTO	FECHA_VENTA
Blue	2002-12-22 12:02:05.033
Blues on the Bayou	2002-12-24 14:15:09.673

Ambas filas en los resultados de la consulta tienen un valor DISCO_COMPACTO que inicia con Blue. Ninguna otra fila cumple con la condición de búsqueda especificada en la cláusula WHERE.

NOTA

El manejo de datos de fecha y hora varía considerablemente a través de las implementaciones de SQL, por lo que los resultados de su DBMS pueden variar enormemente con respecto a la columna FECHA_VENTA.

Utilizar las funciones de valor de cadena UPPER y LOWER

Las funciones de valor de cadena UPPER y LOWER son muy similares en que ambas son utilizadas para convertir caracteres entre mayúsculas y minúsculas. La función UPPER permite convertir una cadena de caracteres completa a mayúsculas. La función LOWER permite convertir una cadena de caracteres completa a minúsculas. Por ejemplo, supongamos que se quiere modificar la instrucción SELECT mostrada en el último ejemplo para arrojar todos los valores DISCO_COMPACTO en mayúsculas. Su instrucción SELECT ahora incluiría una función UPPER:

```
SELECT UPPER(DISCO_COMPACTO) AS TITULO, FECHA_VENTA
FROM FECHAS_VENTAS
WHERE SUBSTRING(DISCO_COMPACTO FROM 1 FOR 4) = 'Blue';
```

Los resultados de la consulta son los mismos que en el último ejemplo, sólo que esta vez en los títulos de CD están todos en mayúsculas, como se muestra en los siguientes resultados:

Título	FECHA_VENTA
-----	-----
BLUE	2002-12-22 12:02:05.033
BLUES ON THE BAYOU	2002-12-24 14:15:09.673

Si se hubiera utilizado la función LOWER en lugar de UPPER, los títulos de CD estarían todos en minúsculas, incluso sin mayúscula al inicio de las palabras. Estas funciones también son muy útiles para comparar datos en implementaciones que son sensibles a mayúsculas y minúsculas cuando no se sabe cuál de las dos fue utilizada para almacenar los datos, o cuando se quiere estar seguro de que los datos aparezcan en mayúsculas o minúsculas al momento que se están insertando, actualizando o convirtiendo datos de una base de datos a otra.

Trabajar con funciones de valor de fecha y hora

Las funciones de valor de fecha y hora proporcionan información acerca de la fecha y la hora actuales. Cada función arroja un valor basado en la fecha u hora (o ambos) tal como están configurados en el sistema operativo. SQL:2006 soporta cinco funciones de valor de fecha y hora, que se describen en la tabla 10-1.

NOTA

Las implementaciones SQL varían ampliamente en cuanto a cómo implementan la funcionalidad fecha y hora; consecuentemente, la implementación de funciones fecha y hora también varía. Por ejemplo, SQL Server soporta solamente la función de valor fecha y hora CURRENT_TIMESTAMP. Por otro lado, Oracle soporta las funciones de valor fecha y hora CURRENT_DATE, CURRENT_TIMESTAMP y LOCALTIMESTAMP, pero no las funciones CURRENT_TIME y LOCALTIME. No obstante, MySQL soporta todas las cinco. Adicionalmente, los valores exactos generados por estas funciones también pueden variar de implementación a implementación. Por ejemplo, los resultados de consulta no siempre incluirán información acerca del huso horario actual, y algunos representan la hora utilizando un formato de 24 horas en lugar de A.M. y P.M.

Función de valor	Descripción
CURRENT_DATE	Arroja un valor que representa la fecha actual.
CURRENT_TIME	Arroja un valor que representa la hora actual. El valor incluye información acerca del huso horario actual, concerniente a Coordinated Universal Time (UCT), o lo que antes era llamado Greenwich Mean Time (GMT).
CURRENT_TIMESTAMP	Arroja un valor que representa la fecha y la hora actuales. El valor incluye información acerca del huso horario actual, concerniente al UCT.
LOCALTIME	Arroja un valor que representa la hora actual.
LOCALTIMESTAMP	Arroja un valor que representa la fecha y la hora actuales.

Tabla 10-1 Funciones de valor fecha y hora soportadas por SQL:2006.

Debido a que la función de valor fecha y hora `CURRENT_TIMESTAMP` es soportada tanto por SQL Server como por Oracle, démosle un vistazo más detallado. No obstante, tenga en mente que implementar cualquiera de las funciones fecha y hora de SQL es un proceso similar, dependiendo de cuáles funciones se soporten por la implementación específica SQL en la que se esté trabajando. Al comprender cómo trabaja la función `CURRENT_TIMESTAMP`, se obtendrá una mucho mejor comprensión de cómo trabajan todas las funciones. Sin embargo, asegúrese de revisar la documentación de su implementación para mayor información sobre cualquiera de las funciones que sean soportadas por ese producto.

Dependiendo de la implementación SQL, se puede utilizar la función `CURRENT_TIMESTAMP` en una instrucción `SELECT` para simplemente recuperar la información actual de la marca de fecha y hora. Como se puede esperar de cualquier cuestión relacionada con la funcionalidad fecha y hora, la forma en que se convoca una función puede variar. Sin embargo, en algunos casos es posible utilizar una instrucción tan sencilla como la siguiente:

```
SELECT CURRENT_TIMESTAMP
```

Esta instrucción recuperará la fecha y hora actuales en algunas implementaciones. En otras implementaciones, puede ser necesario agregar una cláusula `FROM` a la instrucción para recuperar esta información. Por ejemplo, Oracle proporciona una tabla simulada llamada específicamente `DUAL` porque requiere una cláusula `FROM` en todas las instrucciones `SELECT`. Sin importar cómo se necesite escribir la instrucción `SELECT`, con toda probabilidad no será muy aprovechable utilizar una función `CURRENT_TIMESTAMP` de esta manera. Probablemente se hará un mejor uso de las funciones fecha y hora utilizándolas para comparar datos o para insertar datos automáticamente.

NOTA

La mayoría de las implementaciones SQL tienen funciones especiales para manejar los datos de fecha y hora. Por ejemplo, SQL Server proporciona la función `getdate` para arrojar la fecha actual, mientras que Oracle proporciona el valor especial `SYSDATE` para el mismo propósito. Las primeras implementaciones SQL no incluían ningún soporte para los tipos de datos fecha y hora, pero tan pronto se volvieron populares las bases de datos relacionales en aplicaciones de negocios, los usuarios empezaron a pedirlos. Esto empezó a confundir a los fabricantes al agregar nuevas características, y debido a que no había un estándar SQL a seguir, el resultado fue una amplia variación entre las implementaciones. Productos como MySQL que fueron desarrollados después del estándar tienen muy pocas de esas variaciones. Como siempre, revise la información del fabricante para mayores detalles.

Por ejemplo, supongamos que se quiere utilizar la tabla `FECHAS_VENTAS` (mostrada en la figura 10-2) para insertar la hora y fecha actuales automáticamente en la tabla cada vez que se agregue una nueva fila. La definición de dicha tabla podría lucir de la siguiente manera:

```
CREATE TABLE FECHAS_VENTAS
( DISCO_COMPACTO VARCHAR(60),
  FECHA_VENTA          DATETIME          DEFAULT CURRENT_TIMESTAMP );
```

En esta definición de tabla se ha asignado un valor por defecto a la columna `FECHA_VENTA` que está basado en la función `CURRENT_TIMESTAMP`. Cada vez que una fila sea agregada a la tabla, el valor de fecha y hora será insertado en la columna `DATE_SOLD` para esa fila. Como

resultado, se pueden crear instrucciones INSERT que especifiquen solamente el valor DISCO_COMPACTO. La fecha y la hora actuales, por lo tanto, son agregadas automáticamente a la columna DATE_SOLD al momento que se agrega la fila.

Utilizar expresiones de valor

Una expresión de valor es un tipo de expresión que arroja un valor de datos. La expresión puede incluir nombres de columna, valores, operadores matemáticos, palabras clave u otros elementos que juntos creen algún tipo de fórmula o expresión que arroje un valor único. Por ejemplo, se pueden combinar los valores de dos columnas para crear un valor, o se pueden realizar operaciones en el valor de una columna para crear un nuevo valor.

En esta sección daremos un vistazo a las expresiones de valor numéricas y también a las expresiones de valor CASE y CAST. Para demostrar cómo funcionan varias de estas expresiones, utilizaremos la tabla RASTREO_CD, mostrada en la figura 10-3.

Trabajar con expresiones de valor numéricas

Las expresiones de valor numéricas son expresiones que utilizan operadores matemáticos para realizar cálculos sobre valores de datos numéricos almacenados en tablas. Se pueden utilizar estos operadores para sumar, sustraer, multiplicar y dividir estos valores. La tabla 10-2 muestra los cuatro operadores que se pueden utilizar para crear expresiones de valor numéricas.

Es posible construir expresiones de valor numéricas de una forma muy parecida a como se construyen las fórmulas matemáticas. Los principios básicos son los mismos. Por ejemplo, la multiplicación y la división toman precedencia sobre la suma y la resta, y los elementos que deben ser

NOMBRE_CD: VARCHAR(60)	CATEGORÍA_CD CHAR(4)	EN_EXISTENCIA: INT	EN_PEDIDO: INT	VENDIDOS: INT
Famous Blue Raincoat	FROK	19	16	34
Blue	CPOP	28	22	56
Court and Spark	CPOP	12	11	48
Past Light	NEWA	6	7	22
That Christmas Feeling	XMAS	14	14	34
Patsy Cline: 12 Greatest Hits	CTRY	15	18	54
Out of Africa	STRK	8	5	26
Leonard Cohen The Best of	FROK	6	8	18
Fundamental	BLUS	10	6	21
Blues on the Bayou	BLUS	11	10	17

Figura 10-3 Utilizar expresiones de valor en la tabla RASTREO_CD.

Expresión	Operador	Ejemplo
Suma	+	EN_EXISTENCIA + EN_PEDIDO
Sustracción	-	VENDIDOS - (EN_EXISTENCIA + EN_ORDEN)
Multiplicación	*	EN_EXISTENCIA * 2
División	/	VENDIDOS / 2

Tabla 10-2 Utilizar expresiones de valor numéricas para calcular datos.

calculados primero se encierran en paréntesis; de otra manera, cada operación se calcula de acuerdo con la precedencia y orden en la cual fue descrita. Por ejemplo, la fórmula $2 + 2 * 5 / 4$ es igual a 4.5; sin embargo, la fórmula $(2 + 2) * 5 / 4$ es igual a 5. En la primer fórmula, el 2 fue multiplicado por 5, luego dividido entre 4 y luego se le sumaron 2. En la segunda fórmula, al 2 se le sumaron 2, luego fue multiplicado por 5 y luego dividido entre 4.

Ahora demos un vistazo a un ejemplo de una expresión de valor numérica. Supongamos que se quiere agregar la columna EN_EXISTENCIA a la columna EN_PEDIDO en la tabla RASTREO_CD. Se puede crear una instrucción SELECT similar a la siguiente:

```
SELECT NOMBRE_CD, EN_EXISTENCIA, EN_PEDIDO, (EN_EXISTENCIA + EN_PEDIDO)
AS TOTAL
FROM RASTREO_CD
```

Como se puede ver, la cláusula SELECT especifica primero tres nombres de columna: NOMBRE_CD, EN_EXISTENCIA y EN_PEDIDO. Por lo regular éstos son seguidos por una expresión de valor numérica: (EN_EXISTENCIA + EN_PEDIDO). Los valores de las columnas EN_EXISTENCIA y EN_PEDIDO se agregan juntos y se incluyen en los resultados de la consulta bajo una columna llamada TOTAL, como se muestra en los siguientes resultados:

NOMBRE_CD	EN_EXISTENCIA	EN_PEDIDO	TOTAL
-----	-----	-----	-----
Famous Blue Raincoat	19	16	35
Blue	28	22	50
Court and Spark	12	11	23
Past Light	6	7	13
That Christmas Feeling	14	14	28
Patsy Cline: 12 Greatest Hits	15	18	33
Out of Africa	8	5	13
Leonard Cohen The Best Of	6	8	14
Fundamental	10	6	16
Blues on the Bayou	11	10	21

Para cada fila se ha agregado un valor a la columna TOTAL que junta los valores en la columna EN_EXISTENCIA y en la columna EN_PEDIDO.

Las expresiones de valor numéricas no están limitadas a la cláusula SELECT. Por ejemplo, se puede utilizar una en una cláusula WHERE para especificar una condición de búsqueda. Supongamos que se quieren arrojar los mismos resultados que en la instrucción SELECT anterior pero sólo

para aquellos CD con un valor TOTAL mayor a 25. Se puede modificar la instrucción de la manera siguiente:

```
SELECT NOMBRE_CD, EN_EXISTENCIA, EN_PEDIDO, (EN_EXISTENCIA + EN_PEDIDO)
AS TOTAL
FROM RASTREO_CD
WHERE (EN_EXISTENCIA + EN_PEDIDO) > 25
```

Ahora los resultados de la búsqueda incluirán solamente 4 filas, como se muestra a continuación:

NOMBRE_CD	EN_EXISTENCIA	EN_PEDIDO	TOTAL
Famous Blue Raincoat	19	16	35
Blue	28	22	50
That Christmas Feeling	14	14	28
Patsy Cline: 12 Greatest Hits	15	18	33

Los operadores de valor numéricos pueden también ser combinados entre sí para crear expresiones más complejas. En el siguiente ejemplo se incluye una expresión adicional que calcula tres conjuntos de valores y los combina en una columna en los resultados de la consulta:

```
SELECT NOMBRE_CD, EN_EXISTENCIA, EN_PEDIDO, (EN_EXISTENCIA + EN_PEDIDO)
AS TOTAL,
SOLD, (SOLD - (EN_EXISTENCIA + EN_PEDIDO)) AS ESCASEZ
FROM RASTREO_CD
WHERE (EN_EXISTENCIA + EN_PEDIDO) > 25
```

Esta instrucción permite calcular cuántos CD se tienen disponibles (EN_EXISTENCIA + EN_PEDIDO) y se comparan con cuántos se han vendido. Luego la diferencia se agrega en la columna ESCASEZ en los resultados de la consulta. Si se han vendido más CD de los que están disponibles, un número positivo se coloca en la columna ESCASEZ. Si, por otro lado, hay suficientes CD disponibles, se crea un número negativo. Los siguientes resultados de la consulta muestran las cantidades calculadas cuando se ejecuta la instrucción SELECT:

NOMBRE_CD	EN_EXISTENCIA	EN_PEDIDO	TOTAL	VENDIDOS	ESCASEZ
Famous Blue Raincoat	19	16	35	34	-1
Blue	28	22	50	56	6
That Christmas Feeling	14	14	28	34	6
Patsy Cline: 12 Greatest Hits	15	18	33	54	21

Los resultados de la consulta ahora incluyen dos columnas calculadas: TOTAL y ESCASEZ. Todos los demás valores (EN_EXISTENCIA, EN_PEDIDO y VENDIDOS) se toman directamente de la tabla.

Como se puede ver, las expresiones de valor numéricas son muy flexibles y pueden ser utilizadas de muchas diferentes formas. Como un agregado a los métodos que se han visto hasta ahora, se pueden combinar valores de columna con valores específicos. Por ejemplo, supongamos que se

quiere revisar cuántos CD se tendrían disponibles si se duplicara la cantidad que se tenía bajo pedido para aquellos CD en los que hay menos de 15 disponibles:

```
SELECT NOMBRE_CD, EN_EXISTENCIA, EN_PEDIDO, (EN_EXISTENCIA + EN_PEDIDO)
AS TOTAL,
      (EN_EXISTENCIA + EN_PEDIDO * 2) AS DOBLE_ORDEN
FROM RASTREO_CD
WHERE (EN_EXISTENCIA + EN_PEDIDO) < 15
```

La segunda expresión de valor numérica en esta instrucción multiplica el valor EN_PEDIDO por 2, lo suma al valor EN_EXISTENCIA, e inserta el total en la columna DOBLE_ORDEN de los resultados de la consulta, como se muestra en los siguientes resultados:

NOMBRE_CD	EN_EXISTENCIA	EN_PEDIDO	TOTAL	DOBLE_ORDEN
-----	-----	-----	-----	-----
Past Light	6	7	13	20
Out of Africa	8	5	13	18
Leonard Cohen The Best Of	6	8	14	22

Los resultados de la consulta incluyen solamente tres filas que cumplen la condición de la cláusula WHERE. Para estas filas, las columnas EN_EXISTENCIA y EN_PEDIDO se calculan para proporcionar los datos que puedan ser útiles para el usuario, dependiendo de sus necesidades. La mejor parte es que estos valores no tienen que ser almacenados en la base de datos. En su lugar, son calculados cuando se ejecuta la instrucción SELECT, en lugar de tener que mantener tablas con datos adicionales.

Utilizar la expresión de valor CASE

Una expresión de valor CASE permite determinar una serie de condiciones que modifican valores específicos arrojados por la instrucción SQL. Se puede cambiar la forma en que un valor está representado o se calcula un nuevo valor. Cada valor es modificado de acuerdo con la condición especificada dentro de la expresión CASE. Una expresión de valor incluye la palabra clave CASE y una lista de condiciones. La última condición proporciona una condición por defecto si ninguna de las condiciones anteriores ha sido cumplida. Entonces la expresión de valor se cierra utilizando la palabra clave END.

Demos un vistazo a un ejemplo para proporcionar una mejor idea de cómo funciona esto. Supongamos que se quiere incrementar el número de CD que se tienen bajo pedido, pero se quiere incrementar la cantidad solamente para ciertos CD. Además, se quiere fundamentar cuántos CD se agregan al pedido sobre la cantidad actual. Antes de actualizar realmente la tabla, se puede ver lo que serían los nuevos valores al crear una instrucción SELECT que consulte la tabla RASTREO_CD, como se muestra en el siguiente ejemplo:

```
SELECT NOMBRE_CD, EN_PEDIDO, NUEVAS_ORDENES =
CASE
  WHEN EN_PEDIDO < 6 THEN EN_PEDIDO + 4
  WHEN EN_PEDIDO BETWEEN 6 AND 8 THEN EN_PEDIDO + 2
  ELSE EN_PEDIDO
END
FROM RASTREO_CD
WHERE EN_PEDIDO < 11;
```

En esta instrucción se especifican tres columnas: NOMBRE_CD, EN_PEDIDO y NUEVAS_ORDENES. La columna NUEVAS_ORDENES es la columna creada por los resultados de la búsqueda. Contendrá los valores actualizados para la expresión de valor CASE. La expresión en sí misma consta del nombre de columna (NUEVAS_ORDENES), el signo igual, la palabra clave CASE, dos cláusulas WHEN/THEN, una cláusula ELSE y la palabra clave END. Cada cláusula WHEN/THEN representa una de las condiciones. Por ejemplo, la primera cláusula especifica que si el valor EN_PEDIDO es menor a 6, entonces 4 deberá ser agregado a ese valor. La segunda cláusula WHEN/THEN especifica que si el valor EN_PEDIDO cae dentro del rango de 6 a 8, entonces 2 deberá ser agregado al valor.

Después de las cláusulas WHEN/THEN, la cláusula ELSE especifica la condición final. Si el valor no cumple las condiciones definidas en las cláusulas WHEN/THEN, la cláusula ELSE especifica una condición por defecto. En el caso de la instrucción SELECT anterior, la cláusula ELSE simplemente se refería a la columna EN_PEDIDO, sin especificar ninguna modificación. (Esto resulta lo mismo que definir EN_PEDIDO + 0.) En otras palabras, si ninguna de las condiciones WHEN/THEN se cumplen, el valor EN_PEDIDO permanece igual. Si se ejecutara la instrucción SELECT, se recibirían los siguientes resultados:

NOMBRE_CD	EN_PEDIDO	NUEVAS_ORDENES
-----	-----	-----
Past Light	7	9
Out of Africa	5	9
Leonard Cohen The Best Of	8	10
Fundamental	6	8
Blues on the Bayou	10	10

Como se puede ver, la fila Out of Africa se incrementa en 4, la fila Blues on the Bayou no cambia en lo absoluto, y las otras tres filas se incrementan en 2.

Además de modificar los valores, se puede utilizar una expresión de valor CASE para renombrar valores. Esto es particularmente útil si los resultados de la consulta incluyen valores que no son reconocibles fácilmente. Por ejemplo, supongamos que se quiere crear una consulta que arroje datos desde la columna CD_CATEGORY de la tabla RASTREO_CD. Se pueden renombrar los valores en la columna para que la información arrojada sea más entendible para los usuarios, como se muestra en la siguiente instrucción SELECT:

```
SELECT NOMBRE_CD, CATEGORIA_CD =
CASE
  WHEN CATEGORIA_CD = 'FROK' THEN 'Folk Rock'
  WHEN CATEGORIA_CD = 'CPOP' THEN 'Classic Pop'
  WHEN CATEGORIA_CD = 'NEWA' THEN 'New Age'
  WHEN CATEGORIA_CD = 'XMAS' THEN 'Christmas'
  WHEN CATEGORIA_CD = 'CTRY' THEN 'Country'
  WHEN CATEGORIA_CD = 'STRK' THEN 'Soundtrack'
  WHEN CATEGORIA_CD = 'BLUS' THEN 'Blues'
  ELSE NULL
END
FROM RASTREO_CD;
```

NOTA

No es necesario colocar los diferentes componentes de la expresión de valor CASE en líneas separadas, como se hizo aquí. Se hizo de esa forma para mostrar más claramente cada componente. También devuelve el código más legible para cualquiera que lo revise.

En esta instrucción SELECT, los diferentes valores en la columna CATEGORIA_CD fueron renombrados a nombres más comunes. Observe que no se necesitan repetir los nombres de columna a la derecha de la palabra clave THEN. La construcción del predicado se asume por el contexto de la cláusula. Cuando se ejecuta esta instrucción, se reciben los siguientes resultados de la consulta:

NOMBRE_CD	CATEGORIA_CD
-----	-----
Famous Blue Raincoat	Folk Rock
Blue	Classic Pop
Court and Spark	Classic Pop
Past Light	New Age
That Christmas Feeling	Christmas
Patsy Cline: 12 Greatest Hits	Country
Out of Africa	Soundtrack
Leonard Cohen The Best Of	Folk Rock
Fundamental	Blues
Blues on the Bayou	Blues

Como se puede ver, nombres más fáciles de utilizar aparecen en la columna CATEGORIA_CD. Si ninguno de los valores originales cumple con la condición definida en las cláusulas WHEN/THEN, un valor nulo será insertado en los resultados de la consulta.

Pregunta al experto

P: ¿Se puede utilizar una expresión de valor CASE en una instrucción diferente a SELECT?

R: También se puede hacer uso de la expresión de valor CASE en la cláusula SET o en una instrucción UPDATE. Por ejemplo, supongamos que se quieren actualizar los valores en la columna EN_PEDIDO en la tabla RASTREO_CD (mostrada en la figura 10-3). Es posible actualizar esos valores definiendo condiciones específicas en una expresión CASE:

```
UPDATE RASTREO_CD
  SET EN_PEDIDO =
CASE
  WHEN EN_PEDIDO < 6 THEN EN_PEDIDO + 4
  WHEN EN_PEDIDO BETWEEN 6 AND 8 THEN EN_PEDIDO + 2
  ELSE EN_PEDIDO
END
```

(continúa)

Esta instrucción agregará 4 a los valores EN_PEDIDO que sean menores a 6, y agregará 2 a los valores EN_PEDIDO que caigan dentro del rango de 6 a 8. De otra manera, no se cambiarán filas adicionales.

P: ¿Se puede hacer referencia a más de una columna en la expresión de valor CASE?

R: Sí, se puede hacer referencia a más de una columna. Supongamos que se quieren actualizar los valores EN_PEDIDO, pero basando esas actualizaciones en los valores CATEGORIA_CD. Se puede crear una instrucción similar a la siguiente:

```
UPDATE RASTREO_CD
SET EN_PEDIDO =
CASE
WHEN CATEGORIA_CD = 'CPOP' THEN EN_PEDIDO * 3
WHEN CATEGORIA_CD = 'BLUS' THEN EN_PEDIDO * 2
ELSE EN_PEDIDO
END
```

En esta instrucción, los valores EN_PEDIDO son multiplicados por 3 cuando los valores CATEGORIA_CD igualan a CPOP, y los valores EN_PEDIDO son multiplicados por 2 cuando los valores CATEGORIA_CD igualan BLUS. De otra manera, no se cambia ningún valor.

P: ¿Existe alguna implicación de rendimiento cuando se utiliza la expresión de valor CASE?

R: Aunque no existen detalles inherentes al rendimiento relacionados al uso de las expresiones de valor CASE, sí es posible pasar los límites. En general, mientras más compleja es la instrucción, especialmente en términos de anidado y de una lógica condicional elaborada, mayores recursos se necesitarán para analizar gramáticamente y ejecutar la instrucción.

Utilizar la expresión de valor CAST

La expresión de valor CAST sirve a un propósito muy diferente al de la expresión CASE. La expresión CAST permite cambiar el tipo de datos de un valor por un valor literal o cuando se recupera ese valor desde la base de datos. Sin embargo, no cambia el tipo de datos de la columna fuente. Esto es particularmente útil cuando se trabaja con lenguajes de programación en los cuales los tipos de datos no coinciden y se necesita utilizar un común denominador para trabajar con el valor.

Para utilizar la expresión de valor CAST, se debe especificar la palabra clave CAST, y, en paréntesis, proporcionar el nombre de columna, la palabra clave AS, y el nuevo tipo de datos, en ese orden. Para ilustrar esto, regresemos a la tabla FECHAS_VENTAS mostrada en la figura 10-2. La tabla incluye la columna DISCO_COMPACTO y la columna FECHA_VENTA. La columna FECHA_VENTA se configura con el tipo de datos TIMESTAMP. Supongamos que se desea cam-

biar los valores de fecha y hora a cadenas de caracteres. Se puede utilizar la expresión CAST en la cláusula SELECT, como se muestra en la siguiente instrucción:

```
SELECT DISCO_COMPACTO, CAST(FECHA_VENTA AS CHAR(25)) AS CHAR_FECHA
FROM FECHAS_VENTAS
WHERE DISCO_COMPACTO LIKE ('%Blue%')
```

Esta instrucción convierte los valores FECHA_VENTA de los valores TIMESTAMP a los valores CHAR. Como se puede ver, todo lo que se necesita hacer es especificar la palabra clave CAST, seguida por los parámetros en paréntesis que identifican la columna fuente y el nuevo tipo de datos, junto con la palabra clave AS. Cuando se ejecuta esta instrucción, se reciben resultados de consulta similares a lo que se vería si no se hubiera utilizado CAST:

DISCO_COMPACTO	CHAR_FECHA
-----	-----
Famous Blue Raincoat	Dec 22 2002 10:58AM
Blue	Dec 22 2002 12:02PM
Blues on the Bayou	Dec 24 2002 2:15PM

Observe que se puede asignar un nombre a la columna que contiene los nuevos resultados de fecha y hora. En este caso, el nombre de la nueva columna es CHAR_FECHA.

NOTA

Se puede encontrar que, dependiendo de la implementación SQL, cuando un valor de fecha y hora es convertido, el formato cambia ligeramente. Por ejemplo, en SQL Server, un valor de fecha es expresado numéricamente y un valor de hora es expresado como un reloj de 24 horas (horario militar), pero cuando valor es convertido a un tipo de datos CHAR, el valor de tiempo es expresado en caracteres alfanuméricos, y la hora se expresa como un reloj de 12 horas (A.M. y P.M).

Utilizar valores especiales

En el capítulo 6 se analizaron valores especiales soportados por SQL que permiten determinar a los usuarios actuales. Un valor especial existe para cada tipo de usuario. Estos valores actúan como marcadores de posición para los valores actuales relacionados con los usuarios. Pueden utilizarse en expresiones para arrojar el valor del usuario específico. SQL soporta cinco valores especiales, que se describen en la tabla 10-3. (Véase el capítulo 6 para mayor información acerca de los diferentes tipos de usuarios SQL.)

Los valores especiales pueden ser utilizados en diferentes formas en una base de datos SQL, por ejemplo para establecer conexiones o ejecutar un procedimiento almacenado. El valor especial, en lugar del nombre del usuario actual, es incrustado en el código para permitir al código permanecer flexible de una situación a la otra. Otra forma en la que un valor especial puede ser utilizado es para almacenar datos de usuario en una tabla. Para ilustrar esto, demos un vistazo a la tabla PEDIDOS_CD en la figura 10-4.

Esta vez se agrega una fila a la tabla, y se inserta un valor para CURRENT_USER dentro de la columna PEDIDO_POR. Esto facilita rastrear cuál usuario ha colocado la orden. Si se observara

Valor	Descripción
CURRENT_USER	Identifica al identificador de usuario actual. Si el identificador de usuario de la sesión SQL es el identificador de usuario actual, entonces CURRENT_USER, USER y SESSION_USER tienen el mismo valor, lo que puede ocurrir si el par de identificador inicial es el único par identificador de usuario activo/nombre de rol (el par en la parte superior de la pila de autenticación).
USER	Identifica al identificador de usuario actual. USER tiene el mismo significado que CURRENT_USER.
SESSION_USER	Identifica al identificador de usuario de la sesión SQL.
CURRENT_ROLE	Identifica el nombre de rol actual.
SYSTEM_USER	Identifica al usuario del sistema operativo actual que invocó un módulo SQL.

Tabla 10-3 Utilizar los valores especiales de SQL:2006.

la definición de la tabla, se vería que un valor por defecto ha sido definido para la columna PEDIDO_POR, como se muestra en la siguiente instrucción CREATE TABLE:

```
CREATE TABLE PEDIDOS_CD
( TITULOCD    VARCHAR(60) ,
  PEDIDO      INT,
  PEDIDO_POR  CHAR(30) DEFAULT CURRENT_USER );
```

Si fueran a insertarse datos a esta tabla, se tendría que especificar solamente un valor TITULO_CD y un valor PEDIDO. El valor PEDIDO_POR sería insertado automáticamente, y ese valor

NOMBRE_CD: VARCHAR(60)	PEDIDO: INT	PEDIDO_POR: CHAR(30)
Famous Blue Raincoat	16	Mngr
Blue	22	AsstMngr
Court and Spark	11	Mngr
Past Light	7	AsstMngr
That Christmas Feeling	14	Mngr
Patsy Cline: 12 Greatest Hits	18	AsstMngr
Out of Africa	5	AsstMngr
Leonard Cohen The Best of	8	Mngr
Fundamental	6	Mngr
Blues on the Bayou	10	Mngr

Figura 10-4 Utilizar el valor especial CURRENT_USER en la tabla PEDIDOS_CD.

sería el identificador de usuario actual. Si no se especifica un valor por defecto para la columna PEDIDO_POR, se puede utilizar el valor especial para insertar al usuario. Por ejemplo, la siguiente instrucción INSERT inserta una fila en la tabla PEDIDOS_CD:

```
INSERT INTO PEDIDOS_CD
VALUES ( 'Rhythm Country and Blues', 14, CURRENT_USER );
```

Cuando se ejecuta la instrucción, un valor representando al identificador de usuario actual (por ejemplo Mngn) se inserta en la columna PEDIDO_POR.

Para determinar la extensión para la que se pueden utilizar los valores especiales, deberá revisar la documentación del producto para su implementación SQL. Encontrará que las formas en las que se pueden utilizar estos valores tendrán variaciones de una implementación a otra; sin embargo, una vez que se sienta cómodo utilizando valores especiales en su implementación, encontrará que son una herramienta muy útil, al mismo tiempo que usted se hace más eficiente con la programación SQL.

Pruebe esto 10-1

Utilizar funciones y expresiones de valor

En este capítulo aprendió acerca de muchas de las funciones y expresiones de valor soportadas por SQL. Ahora se ejercitarán esas funciones y expresiones consultando datos desde la base de datos INVENTARIO. Específicamente, se crearán instrucciones SELECT que contengan las funciones COUNT, MIN, SUM, SUBSTRING y UPPER, y aquellas que contengan expresiones de valor numéricas, CASE y CAST. Puede descargar el archivo Try_This_10.txt (en inglés), que contiene las instrucciones SQL utilizadas en este ejercicio.

Paso a paso

1. Abra la aplicación de cliente para su RDBMS y conéctese con la base de datos INVENTARIO.
2. En la primera instrucción se determinará el número de valores NOMBRE_ARTISTA únicos en la tabla ARTISTAS. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT COUNT(DISTINCT NOMBRE_ARTISTA) AS ARTISTAS
FROM ARTISTAS;
```

La consulta deberá arrojar una cuenta de 18.

3. En la siguiente instrucción se determinará el número mínimo de CD en existencia, como se enlistan en la tabla DISCOS_COMPACTOS. Se nombra la columna en los resultados de la consulta MIN_EXISTENCIA. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT MIN(EN_EXISTENCIA) AS MIN_EXISTENCIA
FROM DISCOS_COMPACTOS;
```

Los resultados de la consulta deberán incluir solamente una columna y una fila, y mostrar un valor de 5. Esto significa que 5 es el número menor de CD que se tienen en existencia para cualquier CD.

(continúa)

4. Ahora se determinará el número total de CD en existencia. Sin embargo, esta vez se agruparán estos totales de acuerdo con los valores ID_DISQUERA. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT ID_DISQUERA, SUM(EN_EXISTENCIA) AS TOTAL
FROM DISCOS_COMPACTOS
GROUP BY ID_DISQUERA;
```

La consulta arrojará 10 filas una por cada valor ID_DISQUERA. El valor TOTAL para cada fila representa el número total de CD para ese grupo particular ID_DISQUERA.

5. En los pasos anteriores se utilizaron funciones set al consultar datos desde la base de datos INVENTARIO. Ahora se practicarán un par de funciones de valor. La primera de éstas es SUBSTRING. En la instrucción SELECT se extraerán los datos desde la columna LUGAR_DE_NACIMIENTO en la tabla ARTISTAS. Se necesita extraer ocho caracteres, empezando por el primer carácter en la cadena. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT NOMBRE_ARTISTA,
       SUBSTRING(LUGAR_DE_NACIMIENTO FROM 1 FOR 8) AS LUGAR_NACIMIENTO
FROM ARTISTAS;
```

Los resultados de la consulta deberán arrojar 18 filas e incluir dos columnas: NOMBRE_ARTISTA y LUGAR_NACIMIENTO. La columna LUGAR_NACIMIENTO contiene los valores extraídos, los cuales están basados en la columna LUGAR_DE_NACIMIENTO de la tabla.

6. La siguiente función de valor que se ejercitará es la función UPPER. En esta instrucción SELECT, los nombres de todos los CD se convertirán a mayúsculas. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT UPPER(TITULO_CD) AS NOMBRE_CD
FROM DISCOS_COMPACTOS;
```

Esta instrucción deberá arrojar 15 filas con una sola columna que enlista el nombre de los CD en la tabla DISCOS_COMPACTOS. Todos los títulos de los CD deberán estar en mayúsculas.

7. Ahora nos moveremos a las expresiones de valor numéricas. La siguiente instrucción que se utilizará crea dos columnas en los resultados de la consulta que duplican y triplican los valores en la columna EN_EXISTENCIA de la tabla DISCOS_COMPACTOS. Sin embargo, la instrucción arroja valores solamente para aquellas filas con un valor EN_EXISTENCIA menor a 25. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT TITULO_CD, EN_EXISTENCIA,
       (EN_EXISTENCIA * 2) AS DOBLE, (EN_EXISTENCIA * 3) AS TRIPLE
FROM DISCOS_COMPACTOS
WHERE EN_EXISTENCIA < 25;
```

La instrucción SELECT deberá arrojar nueve filas, y cada una deberá incluir valores EN_EXISTENCIA que han sido multiplicados por 2 y por 3.

8. La siguiente expresión de valor será la expresión CASE. Esta instrucción proporcionará valores EN_EXISTENCIA actualizados a la columna EN_PEDIDO de los resultados de la consulta. Para los valores EN_EXISTENCIA menores a 10, los valores serán duplicados. Para los valores EN_EXISTENCIA que caigan dentro del rango entre 10 y 15, será sumado un 3. Todos los demás valores EN_EXISTENCIA permanecerán iguales. La instrucción opera solamente en aquellas filas cuyo valor EN_EXISTENCIA original sea menor a 20. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT TITULO_CD, EN_EXISTENCIA, EN_PEDIDO =  
CASE  
    WHEN EN_EXISTENCIA < 10 THEN EN_EXISTENCIA * 2  
    WHEN EN_EXISTENCIA BETWEEN 10 AND 15 THEN EN_EXISTENCIA + 3  
    ELSE EN_EXISTENCIA  
END  
FROM DISCOS_COMPACTOS  
WHERE EN_EXISTENCIA < 20;
```

Los resultados de la consulta deberán incluir solamente siete filas, y la columna EN_PEDIDO de los resultados de la consulta deberá contener los valores actualizados.

9. Ahora se ejercitará la expresión de valor CAST. Se consultará la tabla TIPOS_MUSICA, pero se convertirá el tipo de datos de la columna NOMBRE_TIPO en los resultados de la consulta. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT ID_TIPO, CAST(NOMBRE_TIPO AS CHAR(20)) AS CHAR_TIPO  
FROM TIPOS_MUSICA;
```

Los resultados de la consulta deberán arrojar 11 filas. Los resultados deberán incluir una columna CHAR_TIPO que contenga los valores convertidos.

10. Cierre la aplicación de cliente.

Resumen de Pruebe esto

Ahora se deberá sentir mucho más cómodo con las diferentes funciones y expresiones de valor que se revisaron en este capítulo. Recuerde que cada implementación SQL soporta diferentes funciones y expresiones de valor, usualmente muchas más de las que hemos visto aquí. De hecho, en muchos casos las funciones y expresiones de valor que se vieron en este capítulo representan solamente la punta del *iceberg*. Asegúrese de revisar la documentación de su producto para averiguar qué funciones y valores de expresión se soportan y cómo están implementadas. Encontrará que son herramientas muy útiles en una variedad de situaciones y que vale la pena el esfuerzo que puso en este capítulo.

✓ Autoexamen Capítulo 10

1. ¿Qué es una función set?
2. Se está creando una instrucción SELECT que consulta la tabla CDS_ARTISTA. La tabla incluye las columnas NOMBRE_ARTISTA y NOMBRE_CD. Se requiere que la instrucción arroje el número total de filas en la tabla. ¿Cuál función COUNT deberá incluirse en la cláusula SELECT?
 - A COUNT(*)
 - B COUNT(NOMBRE_ARTISTA)
 - C COUNT(NOMBRE_CD)
 - D COUNT(NOMBRE_ARTISTA, NOMBRE_CD)
3. ¿Cuál función set deberá utilizarse para sumar los valores en una columna?
 - A MAX
 - B COUNT
 - C SUM
 - D AVG
4. Las funciones set requieren que los datos estén_____ de alguna manera.
5. ¿Qué son las funciones de valor?
6. Se está utilizando la función SUBSTRING para extraer caracteres de la columna DISCO_COMPACTO de la tabla FECHAS_VENTAS. Se quiere iniciar con el tercer carácter y extraer ocho caracteres. ¿Qué parámetros deberán utilizarse en la función SUBSTRING?
7. Se está utilizando la función LOWER en el valor Past Light de la columna NOMBRE_CD. ¿Qué valor será arrojado?
8. ¿Qué función arroja un valor que represente la fecha y la hora actuales al igual que la información relacionada con UCT?
 - A LOCALTIMESTAMP
 - B CURRENT_DATE
 - C LOCALTIME
 - D CURRENT_TIMESTAMP
9. ¿Cuáles son los cuatro tipos de operadores que se utilizan en una expresión de valor numérica?
10. Se están consultando datos de la tabla RASTREO_CD. Se quiere agregar valores en la columna EN_EXISTENCIA a los valores en la columna EN_PEDIDO. Luego se quiere duplicar los totales de la columna. ¿Cómo se establece la expresión de valor numérica?

- 11.** ¿Cuál expresión de valor se utiliza para establecer una serie de condiciones que modifiquen valores?
- 12.** Se está creando una instrucción SELECT que incluye una expresión de valor CASE. Se requiere que una de las condiciones especifique que cualquier valor EN_PEDIDO mayor a 10 se incremente en 5. ¿Cómo deberá establecerse la cláusula WHEN/THEN?
- 13.** ¿Cuál es la última palabra en una expresión de valor CASE?
- 14.** ¿Qué es la expresión de valor CAST?
- 15.** Se está consultando la columna FECHA_VENTA en la tabla FECHAS_VENTAS. Se requiere convertir los valores a un tipo de datos CHAR(25), y que los datos sean desplegados en la columna CHAR_FECHA en los resultados de la consulta. ¿Cómo se define la expresión de valor CAST?
- 16.** ¿Qué valor especial puede utilizarse para identificar al identificador de usuario de sesión SQL actual?

Capítulo 11

Acceder a múltiples
tablas

Habilidades y conceptos clave

- Realizar operaciones básicas join
 - Unir tablas con nombres de columna compartidos
 - Utilizar el método join de condición
 - Realizar operaciones de unión
-

Un componente importante de cualquier base de datos relacional es la correlación que puede existir entre dos tablas cualesquiera. Esta relación le permite al usuario publicar datos en una tabla con datos en otra tabla. Este tipo de relaciones es particularmente útil cuando se necesita consultar datos relacionados de más de una tabla y se requiere recuperar esos datos de una forma significativa para que las relaciones entre las tablas sean, para todos los propósitos prácticos, invisibles. Un método que soporta SQL:2006 para consultar datos de esta manera es unir las tablas en una instrucción. Una operación *join* es una operación que hace coincidir las filas en una tabla con las filas de manera tal que las columnas de ambas tablas puedan ser colocadas lado a lado en los resultados de la consulta como si éstos vinieran de una sola tabla. SQL define muchos tipos de operaciones join. El tipo de operación que puede utilizarse en cualquier situación dada depende de la implementación SQL (depende de las instrucciones soportadas y cómo pueda ser impactado el rendimiento), cuáles datos se quieren arrojar y cómo han sido definidas las tablas. En este capítulo se analizarán varias operaciones que combinen datos desde múltiples tablas, incluyendo joins y unions, y proporcionando detalles acerca de cómo son implementadas, además de los resultados que se puedan esperar al utilizarlas.

Realizar operaciones básicas join

Uno de los tipos más simples de operaciones *join* para implementar son las operaciones separadas por comas. En este tipo de operación, sólo se necesita proporcionar una lista de tablas (separadas por comas) en la cláusula FROM de la instrucción SELECT. Se puede, por supuesto, cualificar la operación join en la cláusula WHERE (lo cual es necesario para obtener datos significativos desde las tablas), pero no es obligatorio hacerlo. Sin embargo, antes de analizar la cláusula WHERE, demos primero un vistazo a las operaciones join separadas por comas desde sus puntos más básicos.

Supongamos que se quieren desplegar los datos de la tabla INVENTARIO_CD y de la tabla INTERPRETES, mostradas en la figura 11-1. (La figura también incluye la tabla TIPO_INTER, la cual se estará utilizando en la sección “Crear operaciones join con más de dos tablas”). Se pueden visualizar los datos de las tablas INVENTARIO_CD e INTERPRETES consultando cada tabla por separado, o se pueden unir las tablas en una instrucción.

Para unir las dos tablas, se puede crear una instrucción SELECT tan simple como la siguiente:

```
SELECT *  
FROM INVENTARIO_CD, INTERPRETES;
```

INVENTARIO_CD			INTERPRETES			TIPO_INTER	
NOMBRE_CD: VARCHAR(60)	ID_INTER: INT	EN_EXISTENCIA: INT	ID_INTER: INT	NOMBRE_INTER: VARCHAR(60)	ID_TIPO: INT	ID_TIPO: INT	NOMBRE_TIPO: CHAR(20)
Famous Blue Raincoat	102	12	101	Joni Mitchell	10	10	Popular
Blue	101	24	102	Jennifer Warnes	12	11	Blues
Court and Spark	101	17	103	B.B. King	11	12	Folk
Past Light	105	9	104	Bonnie Raitt	10	13	Rock
Fundamental	104	22	105	William Ackerman	15	14	Classical
Blues on the Bayou	103	19	106	Bing Crosby	16	15	New Age
Longing in Their Hearts	104	18	107	Patsy Cline	17	16	Classic Pop
Luck of the Draw	104	25	108	John Barry	18	17	Country
Deuces Wild	103	17	109	Leonard Cohen	12	18	Soundtrack
Nick of Time	104	11					
Both Sides Now	101	13					

Figura 11-1 Uniendo las tablas INVENTARIO_CD, INTERPRETES y TIPO_INTER.

La consulta produce lo que se conoce como una *tabla de producto cartesiano* (nombrada así debido al matemático y filósofo francés René Descartes), que es una lista de cada fila en una tabla, unida con cada una de las filas en la otra tabla, como se muestra (parcialmente) en los siguientes resultados de la consulta:

NOMBRE_CD	ID_INTER	EN_EXISTENCIA	ID_INTER	NOMBRE_INTER	ID_TIPO
-----	-----	-----	-----	-----	-----
Famous Blue Raincoat	102	12	102	Jennifer Warnes	12
Blue	101	24	102	Jennifer Warnes	12
Court and Spark	101	17	102	Jennifer Warnes	12
Past Light	105	9	102	Jennifer Warnes	12
Fundamental	104	22	102	Jennifer Warnes	12
Blues on the Bayou	103	19	102	Jennifer Warnes	12
Longing in Their Hearts	104	18	102	Jennifer Warnes	12
Luck of the Draw	104	25	102	Jennifer Warnes	12
Deuces Wild	103	17	102	Jennifer Warnes	12
Nick of Time	104	11	102	Jennifer Warnes	12
Both Sides Now	101	13	102	Jennifer Warnes	12
Famous Blue Raincoat	102	12	101	Joni Mitchell	10
Blue	101	24	101	Joni Mitchell	10
Court and Spark	101	17	101	Joni Mitchell	10
Past Light	105	9	101	Joni Mitchell	10

En realidad, la instrucción SELECT anterior arrojará muchas más filas de las que se muestran aquí. Estos resultados representan solamente una lista parcial. Debido a que la tabla INVENTARIO_CD contiene 11 filas y la tabla INTERPRETES contiene 9 filas, los resultados completos de la consulta contendrían 99 filas. Demos un vistazo más profundo a esto. La fila Famous Blue

Raincoat en la tabla INVENTARIO_CD ha sido unida con cada fila en la tabla INTERPRETES, lo cual nos da un total de 9 filas. Cada una de las 10 filas restantes en la tabla INVENTARIO_CD se hace coincidir con cada una de las filas en la tabla INTERPRETES de la misma manera. Como resultado, habrá 99 filas ($11 \times 9 = 99$).

Como se puede ver, los resultados de esta consulta no son de gran ayuda. Sin embargo, es posible generar resultados más significativos si se utiliza una cláusula WHERE para crear una operación *equi-join* (también escrita como *equijoin*), que es un tipo de operación join que equipara los valores de una o más columnas en la primera tabla con los valores de una o más columnas correspondientes en la segunda tabla. Como puede uno imaginar por el nombre, el operador de comparación en una operación equi-join es el operador igual a (=). Por ejemplo, se puede cualificar la instrucción SELECT previa, de la siguiente forma:

```
SELECT *
FROM INVENTARIO_CD, INTERPRETES
WHERE INVENTARIO_CD.ID_INTER = INTERPRETES.ID_INTER;
```

Ahora los resultados de la consulta incluirán solamente a aquellas filas en las cuales el valor en la columna ID_INTER de la tabla INVENTARIO_CD coincida con el valor en la columna ID_INTER de la tabla INTERPRETES. Observe que se deben cualificar los nombres de columna agregando los nombres de tabla. Esto debe hacerse cada vez que las columnas de diferentes tablas tengan el mismo nombre. Si se ejecuta esta instrucción, se arrojarán los siguientes resultados de consulta:

NOMBRE_CD	ID_INTER	EN_EXISTENCIA	ID_INTER	NOMBRE_INTER	ID_TIPO
-----	-----	-----	-----	-----	-----
Famous Blue Raincoat	102	12	102	Jennifer Warnes	12
Blue	101	24	101	Joni Mitchell	10
Court and Spark	101	17	101	Joni Mitchell	10
Past Light	105	9	105	William Ackerman	15
Fundamental	104	22	104	Bonnie Raitt	10
Blues on the Bayou	103	19	103	B.B. King	11
Longing in Their Hearts	104	18	104	Bonnie Raitt	10
Luck of the Draw	104	25	104	Bonnie Raitt	10
Deuces Wild	103	17	103	B.B. King	11
Nick of Time	104	11	104	Bonnie Raitt	10
Both Sides Now	101	13	101	Joni Mitchell	10

Los datos arrojados por esta consulta son ahora mucho más significativos. Cada CD coincide con el artista apropiado, y solamente se despliegan 11 filas, en lugar de 99. Sin embargo, incluso estos resultados de la consulta incluyen datos repetitivos (la columna ID_INTER). Además, puede suceder que no sólo se requieran las columnas duplicadas, sino que se requiera desplegar solamente ciertas columnas y quizá cualificar incluso más la condición de la búsqueda.

Modifiquemos la instrucción SELECT sobre la que se ha estado trabajando al especificar los nombres de las columnas en la cláusula SELECT para limitar las columnas arrojadas y agregar otro predicado a la cláusula WHERE y así limitar las filas arrojadas, como se muestra en el siguiente ejemplo:

```
SELECT INVENTARIO_CD.NOMBRE_CD, INTERPRETES.NOMBRE_INTER, INVENTARIO_
CD.EN_EXISTENCIA
FROM INVENTARIO_CD, INTERPRETES
WHERE INVENTARIO_CD.ID_INTER = INTERPRETES.ID_INTER
AND INVENTARIO_CD.EN_EXISTENCIA < 15;
```

En esta instrucción se ha especificado que tres columnas deberán estar incluidas en los resultados de la consulta. Observe que se han cualificado los nombres de las columnas al incluir los nombres de las tablas. Observe también que la cláusula WHERE incluye un predicado adicional, conectado con el primer predicado por medio de la palabra clave AND. Ahora cualquier fila que sea arrojada deberá también tener valores EN_EXISTENCIA menores a 15. Si se ejecuta esta instrucción, se recibirán los siguientes resultados de la consulta:

NOMBRE_CD	NOMBRE_INTER	EN_EXISTENCIA
-----	-----	-----
Famous Blue Raincoat	Jennifer Warnes	12
Both Sides Now	Joni Mitchell	13
Past Light	William Ackerman	9
Nick of Time	Bonnie Raitt	11

Como se puede ver, se han refinado los resultados de la consulta a solamente la información más esencial. Desde luego, se pueden crear todo tipo de consultas, dependiendo de las necesidades, siempre y cuando se sigan los lineamientos básicos para crear una operación join separada por comas:

- La cláusula FROM debe incluir todos los nombres de las tablas.
- La cláusula WHERE deberá definir una condición join, evitando un producto cartesiano.
- Las referencias de columna deberán ser cualificadas cuando los nombres de columna estén compartidos entre las tablas.

Aparte de estos lineamientos, existe la libertad de crear cualquier tipo de instrucción SELECT que sea necesaria para extraer la información que se requiera de las tablas participantes. Y mientras que utilizar la cláusula WHERE para especificar la condición join era la forma original de realizar operaciones join en SQL, posteriormente en este capítulo se verá que existen otras variaciones de la sintaxis utilizando la palabra clave JOIN, que la mayoría de los programadores de SQL prefieren en lugar de la sintaxis original. Pero sin importar qué sintaxis se utilice, siempre tenga en mente que necesita existir algún tipo de conexión lógica entre las tablas. Esta conexión a menudo se aprecia en la forma de una clave externa, pero no resulta obligatorio de esta manera. (Para mayor información acerca de las claves externas, véase el capítulo 4.) Las tablas pueden ser unidas aun si existe o no una clave externa. Y a pesar de que equi-join es el más común, puede encontrarse ocasionalmente que resulta muy útil usar un operador de comparación join diferente, por ejemplo menor que (<) o BETWEEN, en cuyo caso el operador join es llamado *theta-join*.

Utilizar nombres de correlación

Como se estipuló anteriormente, se deben cualificar las referencias de columna agregando nombres de tabla a aquellas columnas que compartan un nombre. Sin embargo, como una política general, siempre es una buena idea cualificar las referencias de columna cuando se unan tablas, sea necesario o no. Esto vuelve mucho más fácil referenciar el código en algún momento posterior si la instrucción es totalmente autodocumentada. Sin embargo, cuando sus consultas se vuelvan mucho más complejas, se puede volver muy tedioso reingresar los nombres de las tablas cada vez que se hace referencia a una columna. Debido a esto, SQL soporta nombres de correlación, o alias, que pueden ser utilizados durante la duración de una instrucción. Un *nombre de correlación* es simple-

mente una versión más corta del nombre de tabla actual que se utiliza para simplificar el código y hacerlo más legible.

Tomemos, por ejemplo, la última instrucción SELECT que se vio. Se puede moldear esta instrucción utilizando nombres de correlación para ambas tablas:

```
SELECT c.NOMBRE_CD, p.NOMBRE_INTER, c.EN_EXISTENCIA
FROM INVENTARIO_CD AS c, INTERPRETES AS p
WHERE c.ID_INTER = p.ID_INTER
AND c.EN_EXISTENCIA < 15;
```

La instrucción SELECT produce exactamente los mismos resultados que la instrucción anterior, excepto en la cláusula FROM. De hecho, se puede utilizar la cláusula FROM para definir los alias que son utilizados en el resto de la instrucción. En este caso, la tabla INVENTARIO_CD se renombra como c, y la tabla INTERPRETES se renombra como p. Como resultado, c y p deben utilizarse en cualquier otra posición dentro de la instrucción SELECT cuando se haga referencia a esas tablas. Una vez que se haya definido el nombre de la correlación, ya no podrá utilizarse el nombre actual de la tabla. Y esto puede resultar muy confuso debido a que se utiliza el alias en la cláusula SELECT, pero no está aún definido sino hasta la cláusula FROM posterior a la cláusula SELECT. Sin embargo, esto toma sentido si recordamos que la cláusula FROM siempre se procesa primero.

Para comprender mejor cómo funciona el proceso de renombrado, recordemos el tema de cómo son procesadas en las instrucciones SELECT. En el capítulo 7 se estableció que la cláusula FROM se procesa en primer lugar y la cláusula SELECT se procesa en el último. Ésta es la razón por la que los nombres de correlación se definen en la cláusula FROM. Una vez que son definidos, todas las demás cláusulas pueden (y deben) utilizar esos alias cuando se definan las referencias de columna. Los nombres de correlación se utilizan durante todo el resto de la instrucción, pero solamente aplican a la instrucción en la cual están definidos. Si se crea una nueva instrucción SELECT, deben redefinirse esos nombres para poder utilizarlos en la nueva instrucción.

Como se pudo ver en la anterior instrucción SELECT, un nombre de correlación es definido inmediatamente después del nombre actual de la tabla. El nuevo nombre sigue a la palabra clave AS. Sin embargo, la palabra clave AS no es obligatoria. En la mayoría de las implementaciones también puede utilizarse la siguiente convención para renombrar las tablas dentro de una consulta:

```
SELECT c.NOMBRE_CD, p.NOMBRE_INTER, c.EN_EXISTENCIA
FROM INVENTARIO_CD c, INTERPRETES p
WHERE c.ID_INTER = p.ID_INTER
AND c.EN_EXISTENCIA < 15;
```

Observe que sólo se especifica el nuevo nombre, sin la palabra clave AS. Esto hace a la instrucción SQL mucho más sencilla. De hecho, algunas implementaciones como Oracle no permiten utilizar la palabra clave AS en lo absoluto, aun cuando ésta es parte del estándar SQL. Una vez más, esta última instrucción SELECT proporcionará los mismos resultados de la consulta que se vieron en los dos ejemplos anteriores. Sólo que la instrucción en sí misma ha sido cambiada.

Crear operaciones join con más de dos tablas

Hasta este punto, los ejemplos que se han visto han unido solamente dos tablas. Sin embargo, se puede utilizar una operación join separada por comas para desplegar datos de más de dos tablas.

Si nos referimos una vez más a la figura 11-1, se verá que la tabla TIPO_INTER está incluida en la ilustración. Es posible, si se desea, unir las tres tablas en una sola instrucción SELECT, como se muestra en el siguiente ejemplo:

```
SELECT c.NOMBRE_CD, p.NOMBRE_INTER, t.NOMBRE_TIPO
FROM INVENTARIO_CD c, INTERPRETES p, TIPO_INTER t
WHERE c.ID_INTER = p.ID_INTER
      AND p.ID_TIPO = t.ID_TIPO
      AND NOMBRE_TIPO = 'Popular';
```

En esta instrucción, la cláusula FROM incluye todas las tres tablas. Adicionalmente, la cláusula WHERE proporciona dos condiciones equi-join: una que traza las columnas ID_INTER y otra que traza las columnas ID_TIPO. Si se ejecuta esta instrucción, se recibirán los siguientes resultados de la consulta:

NOMBRE_CD	NOMBRE_INTER	NOMBRE_TIPO
-----	-----	-----
Blue	Joni Mitchell	Popular
Court and Spark	Joni Mitchell	Popular
Fundamental	Bonnie Raitt	Popular
Longing in Their Hearts	Bonnie Raitt	Popular
Luck of the Draw	Bonnie Raitt	Popular
Nick of Time	Bonnie Raitt	Popular
Both Sides Now	Joni Mitchell	Popular

Observe que la información de las tres tablas está incluida en los resultados: el nombre del CD, el nombre del artista y la categoría del artista. A pesar de que pudiera existir una relación entre la tabla INVENTARIO_CD y la tabla INTERPRETES, al igual que entre la tabla INTERPRETES y la tabla TIPO_INTER, los resultados de la consulta proporcionan un desplegado uniforme que oculta esas relaciones y muestra solamente la información necesaria.

Crear la operación cross join

Además de la operación join separada por comas, SQL soporta otro tipo de operación llamada cross join. Esta operación es prácticamente idéntica a la operación join separada por comas. La única diferencia es que en lugar de separar los nombres de columna con una coma, se utilizan las palabras clave CROSS JOIN. Por ejemplo, tomemos una instrucción que se utilizó anteriormente y modifiquémosla reemplazándola con las palabras clave CROSS JOIN:

```
SELECT c.NOMBRE_CD, p.NOMBRE_INTER, c.EN_EXISTENCIA
FROM INVENTARIO_CD c CROSS JOIN INTERPRETES p
WHERE c.ID_INTER = p.ID_INTER
      AND c.EN_EXISTENCIA < 15;
```

Esta instrucción arroja tres columnas desde dos tablas, y la cláusula WHERE contiene una condición equi-join. Decidir entre utilizar una o la otra puede simplemente ser una cuestión de determinar cuál instrucción es soportada por la implementación SQL, y, si ambas son soportadas, cuál de ellas proporciona un mejor rendimiento. Con toda probabilidad, se convertirá en un asunto de preferencia personal, con pocas ventajas entre una y otra.

Pregunta al experto

P: Si se están uniendo tablas, parece probable que en algunos casos se arrojarán filas duplicadas en los resultados de la consulta, dependiendo de cómo se construya la instrucción SELECT. ¿Cómo pueden evitarse las filas duplicadas?

R: Al igual que con la mayoría de las consultas, es posible generar filas duplicadas. Por ejemplo, la siguiente instrucción arrojará tipos y nombres de artista duplicados:

```
SELECT P.NOMBRE_INTER, T.NOMBRE_TIPO
FROM INVENTARIO_CD C, INTERPRETES P, TIPO_INTER T
WHERE C.ID_INTER = P.ID_INTER
AND P.ID_TIPO = T.ID_TIPO;
```

Para aquellos artistas que tienen más de un CD, los resultados de la consulta contendrán alguna fila para cada uno de esos CD. Sin embargo, al igual que con cualquier otra instrucción SELECT, puede añadirse la palabra clave DISTINCT a la cláusula SELECT, como se muestran el siguiente ejemplo:

```
SELECT DISTINCT P.NOMBRE_INTER, T.NOMBRE_TIPO
FROM INVENTARIO_CD C, INTERPRETES P, TIPO_INTER T
WHERE C.ID_INTER = P.ID_INTER
AND P.ID_TIPO = T.ID_TIPO;
```

Esta instrucción arrojará menos filas que la instrucción anterior (5 comparadas con 11), y ninguna fila estará duplicada. Observe también que se pueden lograr los mismos resultados sin la palabra clave DISTINCT utilizando una cláusula GROUP BY que enliste ambas columnas.

Crear la operación self-join

Otro tipo de operación join que puede crearse es self-join, que puede ser tanto una operación separada por comas como una operación cross join. En una operación self-join se crea una condición join que se refiere a la misma tabla dos veces, esencialmente uniendo la tabla consigo misma. Esto se realiza casi siempre para resolver una relación recursiva, al encontrar otras filas en la misma tabla que están relacionadas con las filas seleccionadas. Por ejemplo, supongamos que se agrega la tabla EMPLEADOS a la base de datos, como se muestra en la figura 11-2. La tabla EMPLEADOS incluye una lista de números de identificación de empleados, nombres de empleados y los números de identificación de los jefes de los empleados, quienes también están enlistados en la tabla. Por ejemplo, el jefe de Mr. Jones (ID_EMP 102) es Ms. Smith (ID_EMP 101).

Para crear una operación self-join en esta tabla se debe crear una operación join que trate a la tabla como si fueran dos tablas separadas con el mismo nombre, mismas columnas y mismos datos:

```
SELECT a.ID_EMP, a.NOMBRE_EMP, b.NOMBRE_EMP AS ADMINISTRADOR
FROM EMPLEADOS a, EMPLEADOS b
WHERE a.ADMIN = b.ID_EMP
ORDER BY a.ID_EMP;
```


ID_EMP: INT	NOMBRE_EMP: VARCHAR(60)	ADMIN: INT
101	Ms. Smith	NULL
102	Mr. Jones	101
103	Mr. Roberts	101
104	Ms. Hanson	103
105	Mr. Fields	102
106	Ms. Lee	102
107	Mr. Carver	103

Figura 11-2 La tabla EMPLEADOS unida por una operación self-join.

En esta instrucción, a cada instancia de la tabla se le da un nombre de correlación. Como resultado, ahora se tiene (en este ejemplo) una tabla a y una tabla b. Se toman los valores ID_EMP y NOMBRE_EMP de la tabla a, pero se toma el valor ADMINISTRADOR de la tabla b. La condición equi-join se define en la cláusula WHERE al igualar el valor ADMIN en la tabla a con el valor ID_EMP en la tabla b. Esto proporciona el vínculo que trata una tabla física como dos tablas lógicas. Cuando se ejecuta esta instrucción, se reciben los siguientes resultados de la consulta:

ID_EMP	NOMBRE_EMP	ADMINISTRADOR
102	Mr. Jones	Ms. Smith
103	Mr. Roberts	Ms. Smith
104	Ms. Hanson	Mr. Roberts
105	Mr. Fields	Mr. Jones
106	Ms. Lee	Mr. Jones
107	Mr. Carver	Mr. Roberts

Los resultados incluyen el número de identificación de empleado y el número de cada empleado, junto con el nombre del jefe del empleado. Como se puede ver, la operación self-join puede ser una herramienta muy útil en casos como éste en donde una tabla haga referencia a sí misma.

Unir tablas con nombres de columna compartidos

SQL proporciona dos métodos para configurar operaciones join que pueden utilizarse cuando se está trabajando con columnas que tienen los mismos nombres. Estos dos métodos (join natural y join de columna nombrada) permiten fácilmente especificar una condición join entre dos tablas cuando una o más columnas son iguales dentro de esas tablas. Para poder utilizar cualquiera de estos dos métodos, las tablas deben cumplir con las siguientes condiciones:

- Las columnas unidas deberán compartir el mismo nombre y tener tipos de datos compatibles.
- Los nombres de las columnas unidas no pueden ser cualificados con nombres de tabla.

Cuando se está utilizando ya sea el método join natural o un join de columna nombrada, cada tabla debe compartir por lo menos una columna en común. Por ejemplo, las tablas TITULOS_EN_EXISTENCIA y COSTOS_TITULO, mostradas en la figura 11-3, tienen dos columnas que son iguales: TITULO_CD y TIPO_CD. Observe que cada conjunto de columnas a unir está configurado con el mismo tipo de datos.

Se puede utilizar tanto un método join natural o un join de columna nombrada para unir estas dos tablas. Se describen cada uno de estos tipos de operaciones join en varias de las siguientes secciones, y se utilizan las tablas en la figura 11-3 para ilustrar cómo funciona cada uno de estos métodos.

NOTA

No todas las implementaciones de SQL soportan métodos join naturales o métodos join de columna nombrada. Por ejemplo, SQL Server no soporta ninguno de estos métodos, MySQL soporta joins naturales pero no los joins de columna nombrada, y Oracle soporta ambos.

Crear el método join natural

El método join natural hace coincidir automáticamente las filas de aquellas columnas con el mismo nombre. No es necesario especificar ningún tipo de condición equi-join para los joins naturales. La implementación SQL determina cuáles columnas tienen los mismos nombres y luego intenta hacerlas coincidir. El inconveniente de hacer esto es que no se puede especificar cuáles columnas son comparadas, aunque sí se puede especificar cuáles columnas son incluidas en los resultados de la consulta.

En el siguiente ejemplo se utiliza un join natural para unir la tabla TITULOS_EN_EXISTENCIA con la tabla COSTOS_TITULO:

```
SELECT TITULO_CD, TIPO_CD, c.MENUDEO
FROM TITULOS_EN_EXISTENCIA s NATURAL JOIN COSTOS_TITULO c
WHERE s.INVENTARIO > 15;
```

TITULOS_EN_EXISTENCIA

TITULO_CD: VARCHAR(60)	TIPO_CD: CHAR(20)	INVENTARIO: INT
Famous Blue Raincoat	Folk	12
Blue	Popular	24
Past Light	New Age	9
Blues on the Bayou	Blues	19
Luck of the Draw	Popular	25
Deuces Wild	Blues	17
Nick of Time	Popular	11
Both Sides Now	Popular	13

COSTOS_TITULO

TITULO_CD: VARCHAR(60)	TIPO_CD: CHAR(20)	MAYOREO: NUMERIC(5,2)	MENUDEO: NUMERIC(5,2)
Famous Blue Raincoat	Folk	8.00	16.99
Blue	Popular	7.50	15.99
Court and Spark	Popular	7.95	15.99
Past Light	New Age	6.00	14.99
Fundamental	Popular	8.25	16.99
Blues on the Bayou	Blues	7.25	15.99
Longing in their Hearts	Popular	7.50	15.99
Deuces Wild	Blues	7.45	14.99
Nick of Time	Popular	6.95	14.99

Figura 11-3 Unir las tablas TITULOS_EN_EXISTENCIA y COSTOS_TITULO.

En esta instrucción, las tablas están unidas mediante las columnas TITULO_CD y TIPO_CD. Observe que ningún nombre de columna está cualificado (los nombres cualificados no se permiten en operaciones join naturales). Si cualquiera de estos nombres de columna hubiera sido incluido en la cláusula WHERE, tampoco estaría cualificado de esta manera. Cuando se ejecute esta instrucción, se recibirán los siguientes resultados de consulta:

TITULO_CD	TIPO_CD	MENUDEO
-----	-----	-----
Blues on the Bayou	Blues	15.99
Deuces Wild	Blues	14.99
Blue	Popular	15.99

Como se puede ver, solamente son arrojadas tres filas. Existen filas en las cuales los valores de TITULO_CD en ambas tablas son iguales y los valores de TIPO_CD son iguales. Adicionalmente, los valores INVENTARIO son mayores a 15.

Crear el método join de columna nombrada

A pesar de que las operaciones join naturales pueden ser de mucha utilidad para operaciones join simples, encontrará que no siempre se quiere incluir cada columna coincidente como parte de la condición join. La forma de evitar eso es utilizando una operación *join de columna nombrada*, que permite especificar cuáles columnas coincidentes serán agregadas. Por ejemplo, supongamos que se quiere incluir solamente TITULO_CD en la condición join. Se puede modificar el ejemplo anterior de esta manera:

```
SELECT TITULO_CD, s.TIPO_CD, c.MENUDEO
  FROM TITULOS_EN_EXISTENCIA s JOIN COSTOS_TITULO c
    USING (TITULO_CD)
 WHERE s.Inventario > 15;
```

En esta instrucción se ha removido la palabra clave NATURAL y se ha agregado una cláusula USING, que identifica las columnas coincidentes. Observe que el nombre de columna TIPO_CD ahora ha sido cualificado, pero no así la columna TITULO_CD. Solamente las columnas identificadas en la cláusula USING están sin cualificar. Esta instrucción arroja los mismos resultados que el ejemplo anterior, aunque no necesariamente será siempre el caso, y dependerá de los datos en las tablas. Sin embargo, si se incluyen ambas columnas coincidentes en la cláusula USING, definitivamente se obtendrían los mismos resultados que se vieron en el join natural. Al identificar todas las columnas coincidentes en la cláusula USING, se está realizando la misma función que con un join natural.

Utilizar el método join de condición

Hasta ahora en este capítulo se han visto las operaciones join separadas por coma, las cross joins, joins naturales y los joins de columna nombrada. En las operaciones join separadas por coma y cross joint, la condición equi-join es definida en la cláusula WHERE. En las operaciones join naturales, la condición equi-join es asumida automáticamente en todas las columnas coincidentes. Y en las join de columna nombrada, la condición equi-join está localizada en cualquier columna coincidente definida en la cláusula USING. El *join de condición* realiza un método diferente a cualquiera

de éstos. En un join de condición, la condición equi-join está definida en la cláusula ON, que funciona de manera muy similar a la cláusula WHERE. Sin embargo, a pesar de utilizar la cláusula ON, una condición básica join es similar de muchas maneras a las operaciones join previas que se han visto, excepto que, a diferencia de las join naturales y de las join de columna nombrada, la condición join permite hacer coincidir cualquier columna compatible de una tabla con cualquier otra de otra tabla. Los nombres de columna no necesitan ser iguales. La join de condición es la sintaxis preferida por la mayoría de los programadores SQL debido a su claridad, flexibilidad y amplio soporte entre todas las implementaciones SQL.

Una join de condición puede ser separada en dos tipos de uniones: *inner joins* y *outer joins*. La diferencia entre estas dos uniones es la cantidad de datos arrojados por la consulta. Una inner join arroja solamente aquellas filas que coinciden con la condición equi-join definida en la instrucción SELECT. En otras palabras, la inner join arroja solamente filas coincidentes. Ésta era la join original disponible en SQL, y por lo tanto algunos programadores la llaman “join estándar”, a pesar de que esto es un error debido a que todas las joins presentadas en este capítulo están descritas en el estándar SQL. Una outer join, por otro lado, arroja las filas coincidentes y alguna o todas las filas no coincidentes, dependiendo del tipo de outer join.

NOTA

De acuerdo con el estándar SQL:2006, tanto las operaciones join naturales como las join de columna nombrada soportan joins inner y outer. Sin embargo, esto puede variar entre las diferentes implementaciones SQL, por lo que debe asegurarse de revisar la documentación del producto. De manera predeterminada, una join se procesa como una inner join, a menos que esté específicamente definida como una outer join.

Crear la inner join

Ahora que se tiene una idea general de la condición join, demos un vistazo más de cerca a la inner join. Ésta es la más común de las condiciones join y está especificada al utilizar las palabras clave INNER JOIN. Sin embargo, la palabra clave INNER no es necesaria. Si se utiliza JOIN por sí sola, se asume que es una inner join. Además de la palabra clave JOIN (especificada en la cláusula FROM), también se puede definir una cláusula ON, que se coloca inmediatamente después de la cláusula FROM. Echemos un vistazo a un ejemplo para ver cómo funciona esto.

Supongamos que se quiere unir las tablas TITULO_CDS y ARTISTAS_TITULOS, mostradas en la figura 11-4. En el siguiente ejemplo se ha creado una inner join que está basada en las columnas ID_TITULO en las dos tablas:

```
SELECT t.TITULO, ta.ID_ARTISTA
FROM TITULO_CDS t INNER JOIN ARTISTAS_TITULOS ta
ON t.ID_TITULO = ta.ID_TITULO
WHERE t.TITULO LIKE ('%Blue%');
```

La instrucción utiliza las palabras clave INNER JOIN para unir las tablas TITULO_CDS y ARTISTAS_TITULOS. La condición equi-join se define en la cláusula ON, utilizando la columna TITLE_ID en cada tabla. Observe que los nombres de correlación han sido definidos en ambas tablas. La instrucción SELECT es cualificada adelante por la cláusula WHERE, que arroja solamente aquellas filas que contienen el valor Blue en la columna TITLE de la tabla TITULO_CDS. Cuando se ejecute esta consulta, se recibirán los siguientes resultados de consulta:

TITULO_CDS		ARTISTAS_TITULOS		ARTISTAS_CD	
ID_TITULO: INT	TITULO: VARCHAR(60)	TITLE_ID: INT	ID_ARTISTA: INT	ID_ARTISTA: INT	ARTISTA: VARCHAR(60)
101	Famous Blue Raincoat	101	2001	2001	Jennifer Warnes
102	Blue	102	2002	2002	Joni Mitchell
103	Court and Spark	103	2002	2003	William Ackerman
104	Past Light	104	2003	2004	Kitaro
105	Kojiki	105	2004	2005	Bing Crosby
106	That Christmas Feeling	106	2005	2006	Patsy Cline
107	Patsy Cline: 12 Greatest Hits	107	2006	2007	Jose Carreras
108	Carreras Domingo Pavarotti in Concert	108	2007	2008	Luciano Pavarotti
109	Out of Africa	108	2008	2009	Placido Domingo
110	Leonard Cohen The Best of	108	2009	2010	John Barry
111	Fundamental	109	2010	2011	Leonard Cohen
112	Blues on the Bayou	110	2011	2012	Bonnie Raitt
113	Orlando	111	2012	2013	B.B. King
		112	2013	2014	David Motion
		113	2014	2015	Sally Potter
		113	2015		

Figura 11-4 Unir las tablas TITULO_CDS, TITLES_ARTISTS y CD_ARTISTS.

TITULO	ID_ARTISTA
-----	-----
Famous Blue Raincoat	2001
Blue	2002
Blues on the Bayou	2013

Como se puede ver, los resultados incluyen información de ambas tablas: la columna TITULO de la tabla TITULO_CDS en la columna ID_ARTISTA de la tabla ARTISTAS_TITULOS. A pesar de que esta información puede ser muy útil, podría resultar mejor para algunos usuarios si ellos pudieran ver los nombres reales de los artistas en lugar de números. La forma de lograr esto es incluir una tercera tabla en la unión.

Volvamos al ejemplo anterior y agreguemos una segunda condición join a la tabla ARTISTAS_CD (mostrada en la figura 11-4). En el siguiente ejemplo se agrega la segunda condición inmediatamente después de la cláusula original ON:

```
SELECT t.TITULO, a.ARTISTA
FROM TITULO_CDS t INNER JOIN ARTISTAS_TITULOS ta
    ON t.ID_TITULO = ta.ID_TITULO
    INNER JOIN ARTISTAS_CD a
    ON ta.ID_ARTISTA = a.ID_ARTISTA
WHERE t.TITULO LIKE ('%Blue%');
```

Observe que se repiten las palabras clave INNER JOIN, seguidas del nombre de la tercera tabla, la cual entonces sigue a la otra cláusula ON. En esta cláusula, la condición equi-join se define en las columnas ID_ARTISTA dentro de las tablas ARTISTAS_TITULOS y ARTISTAS_CD. Tenga en mente que no es necesario incluir la palabra clave INNER, ni tampoco las columnas especificadas en la cláusula ON necesitan tener el mismo nombre.

Si se ejecuta esta instrucción, se obtendrán los siguientes resultados de la consulta:

TITULO	ARTISTA
-----	-----
Famous Blue Raincoat	Jennifer Warnes
Blue	Joni Mitchell
Blues on the Bayou	B.B. King

Observe que los nombres de artista ahora están listados en los resultados. Observe también que el hecho de que se hayan utilizado tres tablas para recuperar esta información es invisible a cualquier persona que vea los resultados de la consulta.

Crear la outer join

Como se mencionó anteriormente en este capítulo, una operación outer join arroja todas las filas coincidentes y alguna o todas las filas no coincidentes, dependiendo del tipo de outer join que se cree. SQL soporta tres tipos de outer joins:

- **Left** Arroja todas las filas coincidentes y todas las filas no coincidentes de la tabla de la izquierda (la tabla a la izquierda de la palabra clave JOIN).
- **Right** Arroja todas las filas coincidentes y todas las filas no coincidentes de la tabla de la derecha (la tabla a la derecha de la palabra clave JOIN).
- **Full** Arroja todas las filas coincidentes y todas las filas no coincidentes de ambas tablas.

NOTA

Debido a que es una característica relativamente nueva, pocas implementaciones SQL soportan actualmente las operaciones full outer join (Oracle y SQL Server pueden hacerlo, pero no MySQL).

Una operación outer join sigue la misma sintaxis que una operación inner join, sólo que en lugar de utilizar las palabras clave INNER JOIN (o solamente la palabra clave JOIN), se aplica LEFT OUTER JOIN, RIGHT OUTER JOIN o FULL OUTER JOIN. Observe que la palabra clave OUTER es opcional. Por ejemplo, se puede especificar LEFT JOIN en lugar de LEFT OUTER JOIN.

La mejor forma de ilustrar las diferencias entre los tipos de outer joins es mostrar diferentes ejemplos de resultados de consulta para cada tipo. Para ilustrar las diferencias, se utilizan las tablas INFO_CD y CD_TIPO, mostradas en la figura 11-5.

INFO_CD

TITULO: VARCHAR(60)	ID_TIPO: CHAR(4)	EXISTENCIA: INT
Famous Blue Raincoat	FROK	19
Blue	CPOP	28
Past Light	NEWA	6
Out of Africa	STRK	8
Fundamental	NPOP	10
Blues on the Bayou	BLUS	11

TIPO_CD

ID_TIPO: CHAR(4)	NOMBRE_TIPO: CHAR(20)
FROK	Folk Rock
CPOP	Classic Pop
NEWA	New Age
CTRY	Country
STRK	Soundtrack
BLUS	Blues
JAZZ	Jazz

Figura 11-5 Unir las tablas INFO_CD y TIPO_CD.

En el primer ejemplo, se define una operación inner join en las dos tablas, solamente para mostrar cómo lucirían normalmente los resultados de la consulta:

```
SELECT i.TITULO, t.NOMBRE_TIPO, i.EXISTENCIA
FROM INFO_CD i JOIN TIPO_CD t
ON i.ID_TIPO = t.ID_TIPO;
```

Esta instrucción arroja los siguientes resultados de la consulta:

TITULO	NOMBRE_TIPO	EXISTENCIA
-----	-----	-----
Famous Blue Raincoat	Folk Rock	19
Blue	Classic Pop	28
Past Light	New Age	6
Out of Africa	Soundtrack	8
Blues on the Bayou	Blues	11

En la mayoría de los casos la operación inner join proporcionará toda la información necesaria. Pero supongamos que se quieren incluir las filas no coincidentes de la tabla INFO_CD. En este caso, se deberá crear una operación left outer join, como se muestra en el siguiente ejemplo:

```
SELECT i.TITULO, t.NOMBRE_TIPO, i.EXISTENCIA
FROM INFO_CD i LEFT OUTER JOIN TIPO_CD t
ON i.ID_TIPO = t.ID_TIPO;
```

Observe que se ha reemplazado JOIN (por INNER JOIN) con LEFT OUTER JOIN. Como se mencionó anteriormente, se puede omitir la palabra clave OUTER en la mayoría de las implementaciones. Si se ejecuta esta instrucción, se recibirán los siguientes resultados de la consulta:

TITULO	NOMBRE_TIPO	EXISTENCIA
-----	-----	-----
Famous Blue Raincoat	Folk Rock	19
Blue	Classic Pop	28
Past Light	New Age	6
Out of Africa	Soundtrack	8
Fundamental	NULL	10
Blues on the Bayou	Blues	11

Como podrá haber notado, la fila Fundamental está ahora incluida en los resultados de la consulta. A pesar de que esta fila no incluye columnas coincidentes, está aún incluida en los resultados de la consulta debido a que es parte de la tabla de la izquierda (*left*). Para esta fila, a la columna NOMBRE_TIPO se le asigna un valor nulo debido a que ningún valor lógico puede ser arrojado para esta columna. El valor nulo sirve como un marcador de posición.

También es posible arrojar las filas no coincidentes de la tabla TIPO_CD, que es la tabla a la derecha de la palabra clave JOIN:

```
SELECT i.TITULO, t.NOMBRE_TIPO, i.EXISTENCIA
FROM INFO_CD i RIGHT OUTER JOIN TIPO_CD t
ON i.ID_TIPO = t.ID_TIPO;
```

Esta instrucción es prácticamente la misma que la instrucción anterior, excepto que ahora se ha especificado RIGHT. La instrucción arroja los siguientes resultados de la consulta:

TITULO	NOMBRE_TIPO	EXISTENCIA
-----	-----	-----
Famous Blue Raincoat	Folk Rock	19
Blue	Classic Pop	28
Past Light	New Age	6
NULL	Country	NULL
Out of Africa	Soundtrack	8
Blues on the Bayou	Blues	11
NULL	Jazz	NULL

Esta vez las columnas no coincidentes de la tabla de la *derecha* se incluyen en los resultados, y los valores nulos se muestran para las columnas TITLE y STOCK.

Si se quiere arrojar todas las filas no coincidentes, se necesitaría modificar la instrucción para definir una operación full outer join:

```
SELECT i.TITULO, t.NOMBRE_TIPO, i.EXISTENCIA
FROM INFO_CD i FULL OUTER JOIN TIPO_CD t
ON i.ID_TIPO = t.ID_TIPO;
```


Esta instrucción arrojará los siguientes resultados de la consulta:

TITULO	NOMBRE_TIPO	EXISTENCIA
-----	-----	-----
Famous Blue Raincoat	Folk Rock	19
Blue	Classic Pop	28
Past Light	New Age	6
Out of Africa	Soundtrack	8
Fundamental	NULL	10
Blues on the Bayou	Blues	11
NULL	Jazz	NULL
NULL	Country	NULL

Como se puede ver, todas las filas coincidentes y no coincidentes se incluyen en los resultados de la consulta. Observe que están incluidas todas las seis filas de la tabla CD_INFO y todas las siete filas de la tabla TIPO_CD.

Realizar operaciones de unión

SQL proporciona un método más para combinar datos desde diferentes tablas de una forma que resulta un poco diferente de las operaciones join mostradas anteriormente en este capítulo. El operador UNION es un método que puede utilizarse para combinar los resultados de múltiples instrucciones SELECT en un solo conjunto de resultados, esencialmente uniendo filas de una consulta con filas de otra. A diferencia de las operaciones join, que agregan columnas desde múltiples tablas y las colocan lado a lado, las operaciones de unión agregan filas al final del conjunto de resultados. Para poder utilizar un operador UNION, cada instrucción SELECT debe producir columnas *compatibles con unión*, lo que significa que cada una debe producir el mismo número de columnas, y las columnas correspondientes deben tener tipos de datos compatibles. Por ejemplo, si la primera columna de la instrucción SELECT produce una columna de caracteres, entonces otras instrucciones SELECT combinadas con ella que utilizan el operador UNION deberán tener un tipo de datos de caracteres en la primera columna en lugar de un tipo de datos numérico o de fecha y hora.

Demos un vistazo a un ejemplo para mostrar lo que esto significa. La figura 11-6 muestra dos tablas: la tabla CDS_CONTINUADOS y la tabla CDS_DESCONTINUADOS. Las tablas son prácticamente idénticas en estructura pero sirven para dos propósitos muy diferentes, que deberán ser obvios si observamos los nombres de las tablas.

Supongamos que se quiere combinar los datos en estas dos tablas para que pueda visualizarse la información de las dos tablas. Es posible, desde luego, ejecutar dos instrucciones SELECT separadas, o se pueden combinar esas instrucciones en una sola que combine la información, como se muestra en el siguiente ejemplo:

```
SELECT *
  FROM CDS_CONTINUADOS
UNION
SELECT *
  FROM CDS_DESCONTINUADOS;
```

CDS_CONTINUADOS

NOMBRE_CD: VARCHAR(60)	TIPO_CD: CHAR(4)	EN_EXISTENCIA: INT
Famous Blue Raincoat	FROK	19
Blue	CPOP	28
Past Light	NEWA	6
Out of Africa	STRK	8
Fundamental	NPOP	10
Blues on the Bayou	BLUS	11

CDS_DESCONTINUADOS

NOMBRE_CD: VARCHAR(60)	TIPO_CD: CHAR(4)	EN_EXISTENCIA: INT
Court and Spark	FROK	3
Kojiki	NEWA	2
That Christmas Feeling	XMAS	2
Patsy Cline: 12 Greatest Hits	CTRY	4
Leonard Cohen The Best of	FROK	3
Orlando	STRK	1

Figura 11-6 Unir las tablas CDS_CONTINUADOS y CDS_DESCONTINUADOS.

Como se puede ver, las dos instrucciones SELECT están combinadas utilizando el operador UNION. Si se ejecuta esta instrucción, se recibirán los siguientes resultados:

NOMBRE_CD	TIPO_CD	EN_EXISTENCIA
-----	-----	-----
Blue	CPOP	28
Blues on the Bayou	BLUS	11
Court and Spark	FROK	3
Famous Blue Raincoat	FROK	19
Fundamental	NPOP	10
Kojiki	NEWA	2
Leonard Cohen The Best Of	FROK	3
Orlando	STRK	1
Out of Africa	STRK	8
Past Light	NEWA	6
Patsy Cline: 12 Greatest Hits	CTRY	4
That Christmas Feeling	XMAS	2

Los resultados incluyen 12 filas de datos, 6 filas para cada tabla. Se pueden limitar los resultados incluso más al especificar condiciones de búsqueda en las cláusulas WHERE. También es posible especificar que la búsqueda solamente arroje columnas específicas, como se muestra en la siguiente instrucción:

```
SELECT TIPO_CD
FROM CDS_CONTINUADOS
UNION
SELECT TIPO_CD
FROM CDS_DESCONTINUADOS;
```

Ahora cuando se genere la consulta, solamente serán desplegados los valores de la columna TIPO_CD:

```
TIPO_CD
-----
BLUS
CPOP
CTRY
FROK
NEWA
NPOP
STRK
XMAS
```

Observe que sólo son arrojadas 8 filas en lugar de 12. Esto se debe a que de manera predeterminada se filtran las filas para evitar duplicados. Si se quieren incluir todas las filas en los resultados de la consulta, sin importar la existencia de valores duplicados, se puede agregar la palabra clave ALL después del operador UNION, como se muestra en el siguiente ejemplo:

```
SELECT TIPO_CD
      FROM CDS_CONTINUADOS
UNION ALL
SELECT TIPO_CD
      FROM CDS_DESCONTINUADOS;
```

Esta instrucción arrojará 12 filas en lugar de 8, con varios valores duplicados.

Como se puede ver, el operador UNION solamente resulta útil en casos muy específicos. Si se requiere mayor control sobre los resultados de la consulta, se deberá utilizar uno de los diferentes tipos de operadores join soportados por SQL.

Pregunta al experto

P: ¿Hay algún tipo de operador join que sea comparable a utilizar el operador UNION?

R: SQL:2006 soporta actualmente un operador union join que realiza muchas de las mismas funciones que el operador UNION. El operador union join es similar al operador full outer join, en términos de cómo son consolidados los resultados de la consulta. Sin embargo, el operador full outer join permite especificar (en la cláusula ON) cuáles columnas se harán coincidir, a diferencia del operador union join. Además, el operador union join no ha sido implementado de manera general en los RDBMS, y ha sido desaprobado en el estándar SQL:2006, lo que significa que es un candidato a eliminarse de futuras versiones de SQL. Así que para cualquier propósito práctico, el operador union join no es algo que deba preocuparle.

Adicionalmente, el estándar SQL:2006 soporta los operadores INTERSECT y EXCEPT, que tienen una sintaxis similar a UNION. INTERSECT funciona de forma parecida a UNION, excepto que arroja solamente las filas que aparecen en los resultados de ambas instrucciones SELECT. EXCEPT, por otro lado, arroja solamente filas que aparezcan en los resultados de la primera instrucción SELECT pero no en los resultados de la segunda. SQL Server soporta tanto INTERSECT como EXCEPT. Oracle y MySQL soportan INTERSECT pero utilizan el operador MINUS en lugar de EXCEPT.

Pruebe esto 11-1 Consultar múltiples tablas

En este capítulo se han introducido gran variedad de operaciones join así como el operador UNION, el cual técnicamente no está considerado como una operación join. Ahora tendrá la oportunidad de practicar varias de estas técnicas join para consultar datos desde la base de datos INVENTARIO. Específicamente, consultará algunas de las tablas que están configuradas con relaciones de clave externa, que son el tipo de relaciones que ligan datos de una tabla con datos en otra tabla. Debido a que no se cambiará ninguno de los datos, siéntase libre de intentar diferentes tipos de operaciones join, incluso más allá de las que se vieron en este ejercicio. Puede descargar el archivo Try_This_11.txt (en inglés), que contiene las instrucciones SQL utilizadas en este ejercicio.

Paso a paso

1. Abra la aplicación de cliente para su RDBMS y conéctese con la base de datos INVENTARIO.
2. El primer tipo de operación a realizar será una operación join separada por comas en las tablas ARTISTAS y CDS_ARTISTA. La operación join utilizará la columna ID_ARTISTA para establecer la condición equi-join. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT * FROM ARTISTAS a, CDS_ARTISTA c
WHERE a.ID_ARTISTA = c.ID_ARTISTA;
```

Los resultados de la consulta deberán incluir 19 filas, y también las columnas ID_ARTISTA de ambas tablas al igual que las columnas NOMBRE_ARTISTA, LUGAR_DE_NACIMIENTO e ID_DISCO_COMPACTO.

3. Ahora se modificará la instrucción anterior para que también pueda unir la tabla DISCOS_COMPACTOS. De esa manera, se puede desplegar el nombre real de los CD. Adicionalmente, se especifican los nombres de las columnas que deberán ser arrojadas. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT d.TITULO_CD, a.NOMBRE_ARTISTA, a.LUGAR_DE_NACIMIENTO
FROM ARTISTAS a, CDS_ARTISTA c, DISCOS_COMPACTOS d
WHERE a.ID_ARTISTA = c.ID_ARTISTA
AND d.ID_DISCO_COMPACTO = c.ID_DISCO_COMPACTO;
```

Los resultados de la consulta deberán incluir una vez más 19 filas. Sin embargo, esta vez los resultados desplegarán solamente las columnas TITULO_CD, NOMBRE_ARTISTA y LUGAR_DE_NACIMIENTO.

4. Ahora convirtamos la anterior instrucción SELECT en una operación cross join. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT d.TITULO_CD, a.NOMBRE_ARTISTA, a.LUGAR_DE_NACIMIENTO
FROM ARTISTAS a CROSS JOIN CDS_ARTISTA c CROSS JOIN DISCOS_COMPACTOS d
WHERE a.ID_ARTISTA = c.ID_ARTISTA
AND d.ID_DISCO_COMPACTO = c.ID_DISCO_COMPACTO;
```

Se recibirán los mismos resultados de la consulta que se arrojaron en la anterior instrucción SELECT.

5. El siguiente tipo de instrucción que deberá ejercitar es una condición join. La primer condición será del tipo inner join. En esta instrucción se unirán tres tablas: DISCOS_COMPACTOS, TIPOS_DISCO_COMPACTO y TIPOS_MUSICA. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT d.TITULO_CD, t.NOMBRE_TIPO
FROM DISCOS_COMPACTOS d JOIN TIPOS_DISCO_COMPACTO dt
ON d.ID_DISCO_COMPACTO = dt.ID_DISCO_COMPACTO
JOIN TIPOS_MUSICA t
ON dt.ID_TIPO_MUSICA = t.ID_TIPO;
```

Los resultados de la consulta deberán incluir 24 filas. Solamente las columnas TITULO_CD y NOMBRE_TIPO deberán ser desplegadas.

6. Ahora modifiquemos la anterior instrucción SELECT para crear una operación full outer join en ambas condiciones join. Ingrese y ejecute la siguiente instrucción SQL:

```
SELECT d.TITULO_CD, t.NOMBRE_TIPO
FROM DISCOS_COMPACTOS d FULL JOIN TIPOS_DISCO_COMPACTO dt
ON d.ID_DISCO_COMPACTO = dt.ID_DISCO_COMPACTO
FULL JOIN TIPOS_MUSICA t
ON dt.ID_TIPO_MUSICA = t.ID_TIPO;
```

Los resultados de la consulta deberán ahora incluir 26 filas en lugar de 24. Esto se debe a que la tabla TIPOS_MUSICA incluye dos filas que son no coincidentes: la fila Jazz y la fila Internacional. En otras palabras, ningún CD coincide con ninguno de estos dos tipos de música.

7. Cierre la aplicación cliente.

Resumen de Pruebe esto

En este ejercicio se crearon operaciones join separadas por comas, operaciones cross y operaciones de condición. Las operaciones join de condición incluían inner y outer joins. Como se puede ver, las operaciones join proporcionan gran flexibilidad cuando se consultan datos desde las tablas en la base de datos. Sin embargo, no son la única solución cuando se necesitan datos de más de una tabla. Una subconsulta a menudo proporcionará la misma funcionalidad que una operación join. En el capítulo 12 se analizarán las subconsultas con gran detalle. Como se verá, éstas proporcionan una forma más para acceder a datos desde múltiples tablas.

Libros Universitarios, juegos, Revistas, Cursos, Software, Sistemas Operativos, Antivirus y más ... Gratis para el Conocimiento...!

www.detodoprogramas.com

Visítanos y compruébalo

DETODOPROGRAMAS.COM



Material para los amantes de la Programación Java, C/C++/C#, Visual.Net, SQL, Python, Javascript, Oracle, Algoritmos, CSS, Desarrollo Web, Joomla, jquery, Ajax y Mucho Mas...

www.detodoprogramacion.com

Visitanos





Autoexamen Capítulo 11

1. Se está utilizando una operación join separada por comas para unir dos tablas. La primera tabla contiene cinco filas y la segunda tres filas. ¿Cuántas filas contendrá la tabla de producto cartesiano?
2. ¿Qué constituye una condición equi-join en una cláusula WHERE?
3. ¿Cuál cláusula contiene la condición equi-join en una operación join separada por comas?
4. ¿Qué lineamientos básicos deberán seguirse cuando se crea una operación join separada por comas?
5. Se está creando una operación join sobre dos tablas. Se asignan nombres de correlación para cada una de estas tablas. ¿Cuáles nombres deberán utilizarse en la cláusula SELECT: los nombres de correlación o los nombres reales de las tablas?
6. ¿Qué tipo de operación join es prácticamente idéntica a la operación join separada por comas?
 - A Join de condición
 - B Join natural
 - C Cross join
 - D Join de columna nombrada
7. ¿Cuántas tablas están contenidas en una operación self-join?
8. ¿Qué lineamientos deberán seguirse cuando se crean operaciones join naturales o de columna nombrada?
9. ¿Cuál es la diferencia entre una operación join natural y una de columna nombrada?
10. ¿Qué tipo de operación join contiene una cláusula USING para especificar la condición equi-join?
11. ¿Cuáles son los dos tipos de operaciones join de condición?
12. ¿Cuáles son los tres tipos de operación outer join?
13. ¿Cuál tipo de operación join de condición deberá utilizarse si se quieren arrojar solamente filas coincidentes?
 - A Inner join
 - B Left outer join
 - C Right outer join
 - D Full outer join
14. ¿Qué tipo de operación join de condición arroja todas las filas coincidentes y no coincidentes?
 - A Inner join
 - B Left outer join
 - C Right outer join
 - D Full outer join

- 15.** ¿Qué tipo de operación join contiene una cláusula ON?
- A** Cross join
 - B** Join separada por comas
 - C** Join natural
 - D** Join de condición
- 16.** Un operador _____ permite combinar instrucciones SELECT separadas en una sola instrucción para unir los datos en un solo resultado de consulta.
- 17.** ¿Qué palabra clave puede utilizarse con un operador UNION para arrojar todas las filas en los resultados de la consulta, sin importar si existen valores duplicados?