MySQL Tutorial
Abstract

This is the MySQL Tutorial from the MySQL 8.0 Reference Manual.

For legal information, see the Legal Notices.

For help with using MySQL, please visit the MySQL Forums, where you can discuss your issues with other MySQL users.

Document generated on: 2022-05-12 (revision: 73190)
# Table of Contents

Preface and Legal Notices .......................................................... v
1 Tutorial .................................................................................. 1
2 Connecting to and Disconnecting from the Server ........................................... 3
3 Entering Queries ........................................................................ 5
4 Creating and Using a Database ......................................................... 9
   4.1 Creating and Selecting a Database .................................................. 10
   4.2 Creating a Table ........................................................................ 11
   4.3 Loading Data into a Table ................................................................ 12
   4.4 Retrieving Information from a Table .............................................. 13
      4.4.1 Selecting All Data ................................................................ 13
      4.4.2 Selecting Particular Rows .................................................... 14
      4.4.3 Selecting Particular Columns ............................................... 15
      4.4.4 Sorting Rows ...................................................................... 16
      4.4.5 Date Calculations .............................................................. 17
      4.4.6 Working with NULL Values ............................................... 19
      4.4.7 Pattern Matching .................................................................. 20
      4.4.8 Counting Rows .................................................................... 22
      4.4.9 Using More Than one Table .............................................. 24
5 Getting Information About Databases and Tables ........................................... 27
6 Using mysql in Batch Mode ................................................................ 29
7 Examples of Common Queries .............................................................. 31
   7.1 The Maximum Value for a Column ............................................. 31
   7.2 The Row Holding the Maximum of a Certain Column ................. 32
   7.3 Maximum of Column per Group .................................................. 32
   7.4 The Rows Holding the Group-wise Maximum of a Certain Column .... 32
   7.5 Using User-Defined Variables .................................................... 33
   7.6 Using Foreign Keys .................................................................... 34
   7.7 Searching on Two Keys ................................................................ 36
   7.8 Calculating Visits Per Day ........................................................... 36
   7.9 Using AUTO_INCREMENT ........................................................ 36
8 Using MySQL with Apache .............................................................. 41
Preface and Legal Notices

This is the MySQL Tutorial from the MySQL 8.0 Reference Manual.

**Licensing information—MySQL 8.0.** This product may include third-party software, used under license. If you are using a *Commercial* release of MySQL 8.0, see the MySQL 8.0 Commercial Release License Information User Manual for licensing information, including licensing information relating to third-party software that may be included in this Commercial release. If you are using a *Community* release of MySQL 8.0, see the MySQL 8.0 Community Release License Information User Manual for licensing information, including licensing information relating to third-party software that may be included in this Community release.

**Legal Notices**

Copyright © 1997, 2022, Oracle and/or its affiliates.

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, then the following notice is applicable:

**U.S. GOVERNMENT END USERS:** Oracle programs (including any operating system, integrated software, any programs embedded, installed or activated on delivered hardware, and modifications of such programs) and Oracle computer documentation or other Oracle data delivered to or accessed by U.S. Government end users are "commercial computer software" or "commercial computer software documentation" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, reproduction, duplication, release, display, disclosure, modification, preparation of derivative works, and/or adaptation of i) Oracle programs (including any operating system, integrated software, any programs embedded, installed or activated on delivered hardware, and modifications of such programs), ii) Oracle computer documentation and/or iii) other Oracle data, is subject to the rights and limitations specified in the license contained in the applicable contract. The terms governing the U.S. Government's use of Oracle cloud services are defined by the applicable contract for such services. No other rights are granted to the U.S. Government.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications that may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Inside are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Epyc, and the AMD logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.

This software or hardware and documentation may provide access to or information about content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible
for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services unless otherwise set forth in an applicable agreement between you and Oracle. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services, except as set forth in an applicable agreement between you and Oracle.

This documentation is NOT distributed under a GPL license. Use of this documentation is subject to the following terms:

You may create a printed copy of this documentation solely for your own personal use. Conversion to other formats is allowed as long as the actual content is not altered or edited in any way. You shall not publish or distribute this documentation in any form or on any media, except if you distribute the documentation in a manner similar to how Oracle disseminates it (that is, electronically for download on a Web site with the software) or on a CD-ROM or similar medium, provided however that the documentation is disseminated together with the software on the same medium. Any other use, such as any dissemination of printed copies or use of this documentation, in whole or in part, in another publication, requires the prior written consent from an authorized representative of Oracle. Oracle and/or its affiliates reserve any and all rights to this documentation not expressly granted above.

Documentation Accessibility

For information about Oracle's commitment to accessibility, visit the Oracle Accessibility Program website at
https://www.oracle.com/corporate/accessibility/.

Access to Oracle Support for Accessibility

Oracle customers that have purchased support have access to electronic support through My Oracle Support. For information, visit
https://www.oracle.com/corporate/accessibility/learning-support.html#support-tab.
Chapter 1 Tutorial

This chapter provides a tutorial introduction to MySQL by showing how to use the mysql client program to create and use a simple database. mysql (sometimes referred to as the “terminal monitor” or just “monitor”) is an interactive program that enables you to connect to a MySQL server, run queries, and view the results. mysql may also be used in batch mode: you place your queries in a file beforehand, then tell mysql to execute the contents of the file. Both ways of using mysql are covered here.

To see a list of options provided by mysql, invoke it with the --help option:

```bash
$ mysql --help
```

This chapter assumes that mysql is installed on your machine and that a MySQL server is available to which you can connect. If this is not true, contact your MySQL administrator. (If you are the administrator, you need to consult the relevant portions of this manual, such as MySQL Server Administration.)

This chapter describes the entire process of setting up and using a database. If you are interested only in accessing an existing database, you may want to skip the sections that describe how to create the database and the tables it contains.

Because this chapter is tutorial in nature, many details are necessarily omitted. Consult the relevant sections of the manual for more information on the topics covered here.
Chapter 2 Connecting to and Disconnecting from the Server

To connect to the server, you usually need to provide a MySQL user name when you invoke `mysql` and, most likely, a password. If the server runs on a machine other than the one where you log in, you must also specify a host name. Contact your administrator to find out what connection parameters you should use to connect (that is, what host, user name, and password to use). Once you know the proper parameters, you should be able to connect like this:

```
$> mysql -h host -u user -p
```

*host* and *user* represent the host name where your MySQL server is running and the user name of your MySQL account. Substitute appropriate values for your setup. The `********` represents your password; enter it when `mysql` displays the `Enter password:` prompt.

If that works, you should see some introductory information followed by a `mysql>` prompt:

```
$> mysql -h host -u user -p
```

The `mysql>` prompt tells you that `mysql` is ready for you to enter SQL statements.

If you are logging in on the same machine that MySQL is running on, you can omit the host and simply use the following:

```
$> mysql -u user -p
```

If, when you attempt to log in, you get an error message such as `ERROR 2002 (HY000): Can't connect to local MySQL server through socket '/tmp/mysql.sock' (2)`, it means that the MySQL server daemon (Unix) or service (Windows) is not running. Consult the administrator or see the section of Installing and Upgrading MySQL that is appropriate to your operating system.

For help with other problems often encountered when trying to log in, see Common Errors When Using MySQL Programs.

Some MySQL installations permit users to connect as the anonymous (unnamed) user to the server running on the local host. If this is the case on your machine, you should be able to connect to that server by invoking `mysql` without any options:

```
$> mysql
```

After you have connected successfully, you can disconnect any time by typing `QUIT` (or `\q`) at the `mysql>` prompt:

```
mysql> QUIT
Bye
```

On Unix, you can also disconnect by pressing Control+D.

Most examples in the following sections assume that you are connected to the server. They indicate this by the `mysql>` prompt.
Chapter 3 Entering Queries

Make sure that you are connected to the server, as discussed in the previous section. Doing so does not in itself select any database to work with, but that is okay. At this point, it is more important to find out a little about how to issue queries than to jump right in creating tables, loading data into them, and retrieving data from them. This section describes the basic principles of entering queries, using several queries you can try out to familiarize yourself with how mysql works.

Here is a simple query that asks the server to tell you its version number and the current date. Type it in as shown here following the mysql> prompt and press Enter:

```
mysql> SELECT VERSION(), CURRENT_DATE;
+-----------+--------------+
| VERSION() | CURRENT_DATE |
+-----------+--------------+
| 5.8.0-m17 | 2015-12-21   |
+-----------+--------------+
```

This query illustrates several things about mysql:

- A query normally consists of an SQL statement followed by a semicolon. (There are some exceptions where an semicolon may be omitted. QUIT, mentioned earlier, is one of them. We'll get to others later.)

- When you issue a query, mysql sends it to the server for execution and displays the results, then prints another mysql> prompt to indicate that it is ready for another query.

- mysql displays query output in tabular form (rows and columns). The first row contains labels for the columns. The rows following are the query results. Normally, column labels are the names of the columns you fetch from database tables. If you're retrieving the value of an expression rather than a table column (as in the example just shown), mysql labels the column using the expression itself.

- mysql shows how many rows were returned and how long the query took to execute, which gives you a rough idea of server performance. These values are imprecise because they represent wall clock time (not CPU or machine time), and because they are affected by factors such as server load and network latency. (For brevity, the “rows in set” line is sometimes not shown in the remaining examples in this chapter.)

Keywords may be entered in any lettercase. The following queries are equivalent:

```
mysql> SELECT VERSION(), CURRENT_DATE;
mysql> select version(), current_date;
mysql> SeLeCt vErSiOn(), current_DATE;
```

Here is another query. It demonstrates that you can use mysql as a simple calculator:

```
mysql> SELECT SIN(PI()/4), (4+1)*5;
+------------------+---------+
| SIN(PI()/4)      | (4+1)*5 |
+------------------+---------+
| 0.70710678118655 |      25 |
+------------------+---------+
```

The queries shown thus far have been relatively short, single-line statements. You can even enter multiple statements on a single line. Just end each one with a semicolon:

```
mysql> SELECT VERSION(); SELECT NOW();
+-----------+
| VERSION() |
+-----------+
| 8.0.13    |
```
A query need not be given all on a single line, so lengthy queries that require several lines are not a problem. *mysql* determines where your statement ends by looking for the terminating semicolon, not by looking for the end of the input line. (In other words, *mysql* accepts free-format input: it collects input lines but does not execute them until it sees the semicolon.)

Here is a simple multiple-line statement:

```
mysql> SELECT
    -> USER()
    -> ,
    -> CURRENT_DATE;
```

In this example, notice how the prompt changes from *mysql>* to `->` after you enter the first line of a multiple-line query. This is how *mysql* indicates that it has not yet seen a complete statement and is waiting for the rest. The prompt is your friend, because it provides valuable feedback. If you use that feedback, you can always be aware of what *mysql* is waiting for.

If you decide you do not want to execute a query that you are in the process of entering, cancel it by typing `\c`:

```
mysql> SELECT
    -> USER()
    -> \c
mysql>
```

Here, too, notice the prompt. It switches back to *mysql>* after you type `\c`, providing feedback to indicate that *mysql* is ready for a new query.

The following table shows each of the prompts you may see and summarizes what they mean about the state that *mysql* is in.

<table>
<thead>
<tr>
<th>Prompt</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>mysql&gt;</em></td>
<td>Ready for new query</td>
</tr>
<tr>
<td><code>-&gt;</code></td>
<td>Waiting for next line of multiple-line query</td>
</tr>
<tr>
<td>`'&gt;</td>
<td>Waiting for next line, waiting for completion of a string that began with a single quote (<code>'</code>)</td>
</tr>
<tr>
<td>`&quot;&gt;</td>
<td>Waiting for next line, waiting for completion of a string that began with a double quote (<code>&quot;</code>)</td>
</tr>
<tr>
<td>`;'</td>
<td>Waiting for next line, waiting for completion of an identifier that began with a backtick (``)</td>
</tr>
<tr>
<td>`/*&gt;</td>
<td>Waiting for next line, waiting for completion of a comment that began with <code>/*</code></td>
</tr>
</tbody>
</table>

Multiple-line statements commonly occur by accident when you intend to issue a query on a single line, but forget the terminating semicolon. In this case, *mysql* waits for more input:

```
mysql> SELECT USER()
    ->
```
If this happens to you (you think you’ve entered a statement but the only response is a -> prompt), most likely mysql is waiting for the semicolon. If you don’t notice what the prompt is telling you, you might sit there for a while before realizing what you need to do. Enter a semicolon to complete the statement, and mysql executes it:

```
mysql> SELECT USER()
    -> ;
+---------------+
| USER()        |
+---------------+
| jon@localhost |
```

The '-> and "-> prompts occur during string collection (another way of saying that MySQL is waiting for completion of a string). In MySQL, you can write strings surrounded by either ' or " characters (for example, 'hello' or "goodbye"), and mysql lets you enter strings that span multiple lines. When you see a '-> or "-> prompt, it means that you have entered a line containing a string that begins with a ' or " quote character, but have not yet entered the matching quote that terminates the string. This often indicates that you have inadvertently left out a quote character. For example:

```
mysql> SELECT * FROM my_table WHERE name = 'Smith AND age < 30;
>
```

If you enter this SELECT statement, then press Enter and wait for the result, nothing happens. Instead of wondering why this query takes so long, notice the clue provided by the '-> prompt. It tells you that mysql expects to see the rest of an unterminated string. (Do you see the error in the statement? The string 'Smith is missing the second single quotation mark.)

At this point, what do you do? The simplest thing is to cancel the query. However, you cannot just type \c in this case, because mysql interprets it as part of the string that it is collecting. Instead, enter the closing quote character (so mysql knows you've finished the string), then type \c:

```
mysql> SELECT * FROM my_table WHERE name = 'Smith AND age < 30;
> '
```

The prompt changes back to mysql>, indicating that mysql is ready for a new query.

The `-> prompt is similar to the '-> and "-> prompts, but indicates that you have begun but not completed a backtick-quoted identifier.

It is important to know what the '->, "->, and `-> prompts signify, because if you mistakenly enter an unterminated string, any further lines you type appear to be ignored by mysql—including a line containing QUIT. This can be quite confusing, especially if you do not know that you need to supply the terminating quote before you can cancel the current query.

**Note**

Multiline statements from this point on are written without the secondary (-> or other) prompts, to make it easier to copy and paste the statements to try for yourself.
Chapter 4 Creating and Using a Database

Table of Contents

4.1 Creating and Selecting a Database ................................................................. 10
4.2 Creating a Table .............................................................................................. 11
4.3 Loading Data into a Table ................................................................................ 12
4.4 Retrieving Information from a Table ............................................................... 13
  4.4.1 Selecting All Data .................................................................................... 13
  4.4.2 Selecting Particular Rows ......................................................................... 14
  4.4.3 Selecting Particular Columns ................................................................... 15
  4.4.4 Sorting Rows ............................................................................................ 16
  4.4.5 Date Calculations .................................................................................... 17
  4.4.6 Working with NULL Values ...................................................................... 19
  4.4.7 Pattern Matching ...................................................................................... 20
  4.4.8 Counting Rows ......................................................................................... 22
  4.4.9 Using More Than one Table ...................................................................... 24

Once you know how to enter SQL statements, you are ready to access a database.

Suppose that you have several pets in your home (your menagerie) and you would like to keep track
of various types of information about them. You can do so by creating tables to hold your data and
loading them with the desired information. Then you can answer different sorts of questions about
your animals by retrieving data from the tables. This section shows you how to perform the following
operations:

• Create a database
• Create a table
• Load data into the table
• Retrieve data from the table in various ways
• Use multiple tables

The menagerie database is simple (deliberately), but it is not difficult to think of real-world situations
in which a similar type of database might be used. For example, a database like this could be used by
a farmer to keep track of livestock, or by a veterinarian to keep track of patient records. A menagerie
distribution containing some of the queries and sample data used in the following sections can be
obtained from the MySQL website. It is available in both compressed tar file and Zip formats at https://
dev.mysql.com/doc/.

Use the SHOW statement to find out what databases currently exist on the server:

```sql
mysql> SHOW DATABASES;
+----------+
| Database |
+----------+
    mysql |
    test   |
    tmp    |
+----------+
```

The mysql database describes user access privileges. The test database often is available as a
workspace for users to try things out.

The list of databases displayed by the statement may be different on your machine; SHOW DATABASES
does not show databases that you have no privileges for if you do not have the SHOW DATABASES
privilege. See SHOW DATABASES Statement.
If the test database exists, try to access it:

```sql
mysql> USE test
Database changed
```

`USE`, like `QUIT`, does not require a semicolon. (You can terminate such statements with a semicolon if you like; it does no harm.) The `USE` statement is special in another way, too: it must be given on a single line.

You can use the test database (if you have access to it) for the examples that follow, but anything you create in that database can be removed by anyone else with access to it. For this reason, you should probably ask your MySQL administrator for permission to use a database of your own. Suppose that you want to call yours menagerie. The administrator needs to execute a statement like this:

```sql
mysql> GRANT ALL ON menagerie.* TO 'your_mysql_name'@'your_client_host';
```

where `your_mysql_name` is the MySQL user name assigned to you and `your_client_host` is the host from which you connect to the server.

### 4.1 Creating and Selecting a Database

If the administrator creates your database for you when setting up your permissions, you can begin using it. Otherwise, you need to create it yourself:

```sql
mysql> CREATE DATABASE menagerie;
```

Under Unix, database names are case-sensitive (unlike SQL keywords), so you must always refer to your database as `menagerie`, not as `Menagerie`, `MENAGERIE`, or some other variant. This is also true for table names. (Under Windows, this restriction does not apply, although you must refer to databases and tables using the same lettercase throughout a given query. However, for a variety of reasons, the recommended best practice is always to use the same lettercase that was used when the database was created.)

**Note**

If you get an error such as `ERROR 1044 (42000): Access denied for user 'micah'@'localhost' to database 'menagerie'` when attempting to create a database, this means that your user account does not have the necessary privileges to do so. Discuss this with the administrator or see Access Control and Account Management.

Creating a database does not select it for use; you must do that explicitly. To make menagerie the current database, use this statement:

```sql
mysql> USE menagerie
Database changed
```

Your database needs to be created only once, but you must select it for use each time you begin a mysql session. You can do this by issuing a `USE` statement as shown in the example. Alternatively, you can select the database on the command line when you invoke mysql. Just specify its name after any connection parameters that you might need to provide. For example:

```
$> mysql -h host -u user -p menagerie
Enter password: ********
```

**Important**

`menagerie` in the command just shown is not your password. If you want to supply your password on the command line after the `-p` option, you must do so with no intervening space (for example, as `-p password`, not as `–p password`). However, putting your password on the command line is not recommended, because doing so exposes it to snooping by other users logged in on your machine.
4.2 Creating a Table

Creating the database is the easy part, but at this point it is empty, as `SHOW TABLES` tells you:

```
mysql> SHOW TABLES;
Empty set (0.00 sec)
```

The harder part is deciding what the structure of your database should be: what tables you need and what columns should be in each of them.

You want a table that contains a record for each of your pets. This can be called the `pet` table, and it should contain, as a bare minimum, each animal's name. Because the name by itself is not very interesting, the table should contain other information. For example, if more than one person in your family keeps pets, you might want to list each animal's owner. You might also want to record some basic descriptive information such as species and sex.

How about age? That might be of interest, but it is not a good thing to store in a database. Age changes as time passes, which means you'd have to update your records often. Instead, it is better to store a fixed value such as date of birth. Then, whenever you need age, you can calculate it as the difference between the current date and the birth date. MySQL provides functions for doing date arithmetic, so this is not difficult. Storing birth date rather than age has other advantages, too:

- You can use the database for tasks such as generating reminders for upcoming pet birthdays. (If you think this type of query is somewhat silly, note that it is the same question you might ask in the context of a business database to identify clients to whom you need to send out birthday greetings in the current week or month, for that computer-assisted personal touch.)
- You can calculate age in relation to dates other than the current date. For example, if you store death date in the database, you can easily calculate how old a pet was when it died.

You can probably think of other types of information that would be useful in the `pet` table, but the ones identified so far are sufficient: name, owner, species, sex, birth, and death.

Use a `CREATE TABLE` statement to specify the layout of your table:

```
mysql> CREATE TABLE pet (name VARCHAR(20), owner VARCHAR(20), species VARCHAR(20), sex CHAR(1), birth DATE, death DATE);
```

`VARCHAR` is a good choice for the `name`, `owner`, and `species` columns because the column values vary in length. The lengths in those column definitions need not all be the same, and need not be 20. You can normally pick any length from 1 to 65535, whatever seems most reasonable to you. If you make a poor choice and it turns out later that you need a longer field, MySQL provides an `ALTER TABLE` statement.

Several types of values can be chosen to represent sex in animal records, such as `m` and `f`, or perhaps `male` and `female`. It is simplest to use the single characters `m` and `f`.

The use of the `DATE` data type for the `birth` and `death` columns is a fairly obvious choice.

Once you have created a table, `SHOW TABLES` should produce some output:

```
mysql> SHOW TABLES;
+---------------------+
<table>
<thead>
<tr>
<th>Tables in menagerie</th>
</tr>
</thead>
<tbody>
<tr>
<td>pet</td>
</tr>
</tbody>
</table>
+---------------------+
```
To verify that your table was created the way you expected, use a `DESCRIBE` statement:

```sql
mysql> DESCRIBE pet;
+---------+-------------+------+-----+---------+-------+
| Field   | Type        | Null | Key | Default | Extra |
+---------+-------------+------+-----+---------+-------+
| name    | varchar(20) | YES  |     | NULL    |       |
| owner   | varchar(20) | YES  |     | NULL    |       |
| species | varchar(20) | YES  |     | NULL    |       |
| sex     | char(1)     | YES  |     | NULL    |       |
| birth   | date        | YES  |     | NULL    |       |
| death   | date        | YES  |     | NULL    |       |
+---------+-------------+------+-----+---------+-------+
```

You can use `DESCRIBE` any time, for example, if you forget the names of the columns in your table or what types they have.

For more information about MySQL data types, see [Data Types](#).

### 4.3 Loading Data into a Table

After creating your table, you need to populate it. The `LOAD DATA` and `INSERT` statements are useful for this.

Suppose that your pet records can be described as shown here. (Observe that MySQL expects dates in `YYYY-MM-DD` format; this may differ from what you are used to.)

<table>
<thead>
<tr>
<th>name</th>
<th>owner</th>
<th>species</th>
<th>sex</th>
<th>birth</th>
<th>death</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluffy</td>
<td>Harold</td>
<td>cat</td>
<td>f</td>
<td>1993-02-04</td>
<td></td>
</tr>
<tr>
<td>Claws</td>
<td>Gwen</td>
<td>cat</td>
<td>m</td>
<td>1994-03-17</td>
<td></td>
</tr>
<tr>
<td>Buffy</td>
<td>Harold</td>
<td>dog</td>
<td>f</td>
<td>1989-05-13</td>
<td></td>
</tr>
<tr>
<td>Fang</td>
<td>Benny</td>
<td>dog</td>
<td>m</td>
<td>1990-08-27</td>
<td></td>
</tr>
<tr>
<td>Bowser</td>
<td>Diane</td>
<td>dog</td>
<td>m</td>
<td>1979-08-31</td>
<td>1995-07-29</td>
</tr>
<tr>
<td>Chirpy</td>
<td>Gwen</td>
<td>bird</td>
<td>f</td>
<td>1998-09-11</td>
<td></td>
</tr>
<tr>
<td>Whistler</td>
<td>Gwen</td>
<td>bird</td>
<td></td>
<td>1997-12-09</td>
<td></td>
</tr>
<tr>
<td>Slim</td>
<td>Benny</td>
<td>snake</td>
<td>m</td>
<td>1996-04-29</td>
<td></td>
</tr>
</tbody>
</table>

Because you are beginning with an empty table, an easy way to populate it is to create a text file containing a row for each of your animals, then load the contents of the file into the table with a single statement.

You could create a text file `pet.txt` containing one record per line, with values separated by tabs, and given in the order in which the columns were listed in the `CREATE TABLE` statement. For missing values (such as unknown sexes or death dates for animals that are still living), you can use `NULL` values. To represent these in your text file, use \N (backslash, capital-N). For example, the record for Whistler the bird would look like this (where the whitespace between values is a single tab character):

```
Whistler        Gwen    bird    \N      1997-12-09      \N
```

To load the text file `pet.txt` into the `pet` table, use this statement:

```sql
mysql> LOAD DATA LOCAL INFILE '/path/pet.txt' INTO TABLE pet;
```

If you created the file on Windows with an editor that uses `\r\n` as a line terminator, you should use this statement instead:

```sql
mysql> LOAD DATA LOCAL INFILE '/path/pet.txt' INTO TABLE pet
        LINES TERMINATED BY '\r\n';
```

(On an Apple machine running macOS, you would likely want to use `LINES TERMINATED BY '\r'`.)
Retrieving Information from a Table

You can specify the column value separator and end of line marker explicitly in the LOAD DATA statement if you wish, but the defaults are tab and linefeed. These are sufficient for the statement to read the file pet.txt properly.

If the statement fails, it is likely that your MySQL installation does not have local file capability enabled by default. See Security Considerations for LOAD DATA LOCAL, for information on how to change this.

When you want to add new records one at a time, the INSERT statement is useful. In its simplest form, you supply values for each column, in the order in which the columns were listed in the CREATE TABLE statement. Suppose that Diane gets a new hamster named “Puffball.” You could add a new record using an INSERT statement like this:

```sql
mysql> INSERT INTO pet
    VALUES ('Puffball','Diane','hamster','f','1999-03-30',NULL);
```

String and date values are specified as quoted strings here. Also, with INSERT, you can insert NULL directly to represent a missing value. You do not use \N like you do with LOAD DATA.

From this example, you should be able to see that there would be a lot more typing involved to load your records initially using several INSERT statements rather than a single LOAD DATA statement.

4.4 Retrieving Information from a Table

The SELECT statement is used to pull information from a table. The general form of the statement is:

```
SELECT what_to_select
FROM which_table
WHERE conditions_to_satisfy;
```

what_to_select indicates what you want to see. This can be a list of columns, or * to indicate “all columns.” which_table indicates the table from which you want to retrieve data. The WHERE clause is optional. If it is present, conditions_to_satisfy specifies one or more conditions that rows must satisfy to qualify for retrieval.

4.4.1 Selecting All Data

The simplest form of SELECT retrieves everything from a table:

```sql
mysql> SELECT * FROM pet;
```

<table>
<thead>
<tr>
<th>name</th>
<th>owner</th>
<th>species</th>
<th>sex</th>
<th>birth</th>
<th>death</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluffy</td>
<td>Harold</td>
<td>cat</td>
<td>f</td>
<td>1993-02-04</td>
<td>NULL</td>
</tr>
<tr>
<td>Claws</td>
<td>Gwen</td>
<td>cat</td>
<td>m</td>
<td>1994-03-17</td>
<td>NULL</td>
</tr>
<tr>
<td>Buffy</td>
<td>Harold</td>
<td>dog</td>
<td>f</td>
<td>1989-05-13</td>
<td>NULL</td>
</tr>
<tr>
<td>Fang</td>
<td>Benny</td>
<td>dog</td>
<td>m</td>
<td>1990-08-27</td>
<td>NULL</td>
</tr>
<tr>
<td>Bowser</td>
<td>Diane</td>
<td>dog</td>
<td>m</td>
<td>1979-08-31</td>
<td>1995-07-29</td>
</tr>
<tr>
<td>Chirpy</td>
<td>Gwen</td>
<td>bird</td>
<td>f</td>
<td>1998-09-11</td>
<td>NULL</td>
</tr>
<tr>
<td>Whistler</td>
<td>Gwen</td>
<td>bird</td>
<td>NULL</td>
<td>1997-12-09</td>
<td>NULL</td>
</tr>
<tr>
<td>Slim</td>
<td>Benny</td>
<td>snake</td>
<td>m</td>
<td>1996-04-29</td>
<td>NULL</td>
</tr>
<tr>
<td>Puffball</td>
<td>Diane</td>
<td>hamster</td>
<td>f</td>
<td>1999-03-30</td>
<td>NULL</td>
</tr>
</tbody>
</table>

This form of SELECT uses *, which is shorthand for “select all columns.” This is useful if you want to review your entire table, for example, after you’ve just loaded it with your initial data set. For example, you may happen to think that the birth date for Bowser doesn’t seem quite right. Consulting your original pedigree papers, you find that the correct birth year should be 1989, not 1979.

There are at least two ways to fix this:

- **Edit the file pet.txt to correct the error, then empty the table and reload it using DELETE and LOAD DATA:**

  ```sql
  mysql> DELETE FROM pet;
  ```
4.4.2 Selecting Particular Rows

As shown in the preceding section, it is easy to retrieve an entire table. Just omit the \texttt{WHERE} clause from the \texttt{SELECT} statement. But typically you don't want to see the entire table, particularly when it becomes large. Instead, you're usually more interested in answering a particular question, in which case you specify some constraints on the information you want. Let's look at some selection queries in terms of questions about your pets that they answer.

You can select only particular rows from your table. For example, if you want to verify the change that you made to Bowser's birth date, select Bowser's record like this:

\begin{code}
mysql> \texttt{SELECT * FROM pet WHERE name = 'Bowser';}
\end{code}

\begin{table}
\begin{tabular}{|c|c|c|c|c|c|}
\hline
name & owner & species & sex & birth & death \\
\hline
Bowser & Diane & dog & m & 1989-08-31 & 1995-07-29 \\
\hline
\end{tabular}
\end{table}

The output confirms that the year is correctly recorded as 1989, not 1979.

String comparisons normally are case-insensitive, so you can specify the name as \texttt{'bowser'}, \texttt{'BOWSER'}, and so forth. The query result is the same.

You can specify conditions on any column, not just \texttt{name}. For example, if you want to know which animals were born during or after 1998, test the \texttt{birth} column:

\begin{code}
mysql> \texttt{SELECT * FROM pet WHERE birth >= '1998-1-1';}
\end{code}

\begin{table}
\begin{tabular}{|c|c|c|c|c|c|}
\hline
name & owner & species & sex & birth & death \\
\hline
Chirpy & Gwen & bird & f & 1998-09-11 & NULL \\
Puffball & Diane & hamster & f & 1999-03-30 & NULL \\
\hline
\end{tabular}
\end{table}

You can combine conditions, for example, to locate female dogs:

\begin{code}
mysql> \texttt{SELECT * FROM pet WHERE species = 'dog' AND sex = 'f';}
\end{code}

\begin{table}
\begin{tabular}{|c|c|c|c|c|c|}
\hline
name & owner & species & sex & birth & death \\
\hline
Buffy & Harold & dog & f & 1989-05-13 & NULL \\
\hline
\end{tabular}
\end{table}

The preceding query uses the \texttt{AND} logical operator. There is also an \texttt{OR} operator:

\begin{code}
mysql> \texttt{SELECT * FROM pet WHERE species = 'snake' OR species = 'bird';}
\end{code}

\begin{table}
\begin{tabular}{|c|c|c|c|c|c|}
\hline
name & owner & species & sex & birth & death \\
\hline
Chirpy & Gwen & bird & f & 1998-09-11 & NULL \\
Whistler & Gwen & bird & NULL & 1997-12-09 & NULL \\
Slim & Benny & snake & m & 1996-04-29 & NULL \\
\hline
\end{tabular}
\end{table}
AND and OR may be intermixed, although AND has higher precedence than OR. If you use both operators, it is a good idea to use parentheses to indicate explicitly how conditions should be grouped:

```sql
mysql> SELECT * FROM pet WHERE (species = 'cat' AND sex = 'm') OR (species = 'dog' AND sex = 'f');
```

<table>
<thead>
<tr>
<th>name</th>
<th>owner</th>
<th>species</th>
<th>sex</th>
<th>birth</th>
<th>death</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claws</td>
<td>Gwen</td>
<td>cat</td>
<td>m</td>
<td>1994-03-17</td>
<td>NULL</td>
</tr>
<tr>
<td>Buffy</td>
<td>Harold</td>
<td>dog</td>
<td>f</td>
<td>1989-05-13</td>
<td>NULL</td>
</tr>
</tbody>
</table>

### 4.4.3 Selecting Particular Columns

If you do not want to see entire rows from your table, just name the columns in which you are interested, separated by commas. For example, if you want to know when your animals were born, select the name and birth columns:

```sql
mysql> SELECT name, birth FROM pet;
```

<table>
<thead>
<tr>
<th>name</th>
<th>birth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluffy</td>
<td>1993-02-04</td>
</tr>
<tr>
<td>Claws</td>
<td>1994-03-17</td>
</tr>
<tr>
<td>Buffy</td>
<td>1989-05-13</td>
</tr>
<tr>
<td>Fang</td>
<td>1990-08-27</td>
</tr>
<tr>
<td>Bowser</td>
<td>1989-08-31</td>
</tr>
<tr>
<td>Chirpy</td>
<td>1998-09-11</td>
</tr>
<tr>
<td>Whistler</td>
<td>1997-12-09</td>
</tr>
<tr>
<td>Slim</td>
<td>1996-04-29</td>
</tr>
<tr>
<td>Puffball</td>
<td>1999-03-30</td>
</tr>
</tbody>
</table>

To find out who owns pets, use this query:

```sql
mysql> SELECT owner FROM pet;
```

<table>
<thead>
<tr>
<th>owner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harold</td>
</tr>
<tr>
<td>Gwen</td>
</tr>
<tr>
<td>Harold</td>
</tr>
<tr>
<td>Benny</td>
</tr>
<tr>
<td>Diane</td>
</tr>
<tr>
<td>Gwen</td>
</tr>
<tr>
<td>Gwen</td>
</tr>
<tr>
<td>Benny</td>
</tr>
<tr>
<td>Diane</td>
</tr>
</tbody>
</table>

Notice that the query simply retrieves the owner column from each record, and some of them appear more than once. To minimize the output, retrieve each unique output record just once by adding the keyword `DISTINCT`:

```sql
mysql> SELECT DISTINCT owner FROM pet;
```

<table>
<thead>
<tr>
<th>owner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benny</td>
</tr>
<tr>
<td>Diane</td>
</tr>
<tr>
<td>Gwen</td>
</tr>
<tr>
<td>Benny</td>
</tr>
<tr>
<td>Diane</td>
</tr>
</tbody>
</table>

You can use a WHERE clause to combine row selection with column selection. For example, to get birth dates for dogs and cats only, use this query:

```sql
mysql> SELECT name, species, birth FROM pet
```
4.4.4 Sorting Rows

You may have noticed in the preceding examples that the result rows are displayed in no particular order. It is often easier to examine query output when the rows are sorted in some meaningful way. To sort a result, use an `ORDER BY` clause.

Here are animal birthdays, sorted by date:

```sql
mysql> SELECT name, birth FROM pet ORDER BY birth;
```

<table>
<thead>
<tr>
<th>name</th>
<th>birth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffy</td>
<td>1989-05-13</td>
</tr>
<tr>
<td>Bowser</td>
<td>1989-08-31</td>
</tr>
<tr>
<td>Fang</td>
<td>1990-08-27</td>
</tr>
<tr>
<td>Fluffy</td>
<td>1993-02-04</td>
</tr>
<tr>
<td>Claws</td>
<td>1994-03-17</td>
</tr>
<tr>
<td>Slim</td>
<td>1996-04-29</td>
</tr>
<tr>
<td>Whistler</td>
<td>1997-12-09</td>
</tr>
<tr>
<td>Chirpy</td>
<td>1998-09-11</td>
</tr>
<tr>
<td>Puffball</td>
<td>1999-03-30</td>
</tr>
</tbody>
</table>

On character type columns, sorting—like all other comparison operations—is normally performed in a case-insensitive fashion. This means that the order is undefined for columns that are identical except for their case. You can force a case-sensitive sort for a column by using `BINARY` like so: `ORDER BY BINARY col_name`.

The default sort order is ascending, with smallest values first. To sort in reverse (descending) order, add the `DESC` keyword to the name of the column you are sorting by:

```sql
mysql> SELECT name, birth FROM pet ORDER BY birth DESC;
```

<table>
<thead>
<tr>
<th>name</th>
<th>birth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Puffball</td>
<td>1999-03-30</td>
</tr>
<tr>
<td>Chirpy</td>
<td>1998-09-11</td>
</tr>
<tr>
<td>Whistler</td>
<td>1997-12-09</td>
</tr>
<tr>
<td>Slim</td>
<td>1996-04-29</td>
</tr>
<tr>
<td>Claws</td>
<td>1994-03-17</td>
</tr>
<tr>
<td>Fluffy</td>
<td>1993-02-04</td>
</tr>
<tr>
<td>Fang</td>
<td>1990-08-27</td>
</tr>
<tr>
<td>Bowser</td>
<td>1989-08-31</td>
</tr>
<tr>
<td>Buffy</td>
<td>1989-05-13</td>
</tr>
</tbody>
</table>

You can sort on multiple columns, and you can sort different columns in different directions. For example, to sort by type of animal in ascending order, then by birth date within animal type in descending order (youngest animals first), use the following query:

```sql
mysql> SELECT name, species, birth FROM pet
ORDER BY species, birth DESC;
```

<table>
<thead>
<tr>
<th>name</th>
<th>species</th>
<th>birth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chirpy</td>
<td>bird</td>
<td>1998-09-11</td>
</tr>
<tr>
<td>Whistler</td>
<td>bird</td>
<td>1997-12-09</td>
</tr>
</tbody>
</table>
4.4.5 Date Calculations

MySQL provides several functions that you can use to perform calculations on dates, for example, to calculate ages or extract parts of dates.

To determine how many years old each of your pets is, use the `TIMESTAMPDIFF()` function. Its arguments are the unit in which you want the result expressed, and the two dates for which to take the difference. The following query shows, for each pet, the birth date, the current date, and the age in years. An alias (age) is used to make the final output column label more meaningful.

```
mysql> SELECT name, birth, CURDATE(),
           TIMESTAMPDIFF(YEAR, birth, CURDATE()) AS age
       FROM pet;
```

```
+----------+------------+------------+------+
| name     | birth      | CURDATE()  | age  |
+----------+------------+------------+------+
| Fluffy   | 1993-02-04 | 2003-08-19 |   10 |
| Claws    | 1994-03-17 | 2003-08-19 |    9 |
| Buffy    | 1989-05-13 | 2003-08-19 |   14 |
| Fang     | 1990-08-27 | 2003-08-19 |   12 |
| Bowser   | 1989-08-31 | 2003-08-19 |   13 |
| Chirpy   | 1998-09-11 | 2003-08-19 |    4 |
| Whistler | 1997-12-09 | 2003-08-19 |    5 |
| Slim     | 1996-04-29 | 2003-08-19 |    7 |
| Puffball | 1999-03-30 | 2003-08-19 |    4 |
+----------+------------+------------+------+
```

The query works, but the result could be scanned more easily if the rows were presented in some order. This can be done by adding an `ORDER BY name` clause to sort the output by name:

```
mysql> SELECT name, birth, CURDATE(),
           TIMESTAMPDIFF(YEAR, birth, CURDATE()) AS age
       FROM pet ORDER BY name;
```

```
+----------+------------+------------+------+
| name     | birth      | CURDATE()  | age  |
+----------+------------+------------+------+
| Bowser   | 1989-08-31 | 2003-08-19 |   13 |
| Buffy    | 1989-05-13 | 2003-08-19 |   14 |
| Chirpy   | 1998-09-11 | 2003-08-19 |    4 |
| Claws    | 1994-03-17 | 2003-08-19 |    9 |
| Fang     | 1990-08-27 | 2003-08-19 |   12 |
| Whistler | 1997-12-09 | 2003-08-19 |    5 |
| Slim     | 1996-04-29 | 2003-08-19 |    7 |
| Puffball | 1999-03-30 | 2003-08-19 |    4 |
+----------+------------+------------+------+
```

To sort the output by `age` rather than `name`, just use a different `ORDER BY` clause:

```
mysql> SELECT name, birth, CURDATE(),
           TIMESTAMPDIFF(YEAR, birth, CURDATE()) AS age
       FROM pet ORDER BY age;
```

```
+----------+------------+------------+------+
| name     | birth      | CURDATE()  | age  |
+----------+------------+------------+------+
| Chirpy   | 1998-09-11 | 2003-08-19 |    4 |
| Puffball | 1999-03-30 | 2003-08-19 |    4 |
+----------+------------+------------+------+
```
A similar query can be used to determine age at death for animals that have died. You determine which animals these are by checking whether the death value is NULL. Then, for those with non-NULL values, compute the difference between the death and birth values:

```
mysql> SELECT name, birth, death, TIMESTAMPDIFF(YEAR,birth,death) AS age FROM pet WHERE death IS NOT NULL ORDER BY age;
```

The query uses death IS NOT NULL rather than death <> NULL because NULL is a special value that cannot be compared using the usual comparison operators. This is discussed later. See Section 4.4.6, “Working with NULL Values”.

What if you want to know which animals have birthdays next month? For this type of calculation, year and day are irrelevant; you simply want to extract the month part of the birth column. MySQL provides several functions for extracting parts of dates, such as YEAR(), MONTH(), and DAYOFMONTH(). MONTH() is the appropriate function here. To see how it works, run a simple query that displays the value of both birth and MONTH(birth):

```
mysql> SELECT name, birth, MONTH(birth) FROM pet;
```

Finding animals with birthdays in the upcoming month is also simple. Suppose that the current month is April. Then the month value is 4 and you can look for animals born in May (month 5) like this:

```
mysql> SELECT name, birth FROM pet WHERE MONTH(birth) = 5;
```

There is a small complication if the current month is December. You cannot merely add one to the month number (12) and look for animals born in month 13, because there is no such month. Instead, you look for animals born in January (month 1).

You can write the query so that it works no matter what the current month is, so that you do not have to use the number for a particular month. DATE_ADD() enables you to add a time interval to a given date. If you add a month to the value of CURDATE(), then extract the month part with MONTH(), the result produces the month in which to look for birthdays:

```
mysql> SELECT name, birth FROM pet
```
WHERE MONTH(birth) = MONTH(DATE_ADD(CURDATE(), INTERVAL 1 MONTH));

A different way to accomplish the same task is to add 1 to get the next month after the current one after using the modulo function (MOD) to wrap the month value to 0 if it is currently 12:

mysql> SELECT name, birth FROM pet
    WHERE MONTH(birth) = MOD(MONTH(CURDATE()), 12) + 1;

MONTH() returns a number between 1 and 12. And MOD(something, 12) returns a number between 0 and 11. So the addition has to be after the MOD(), otherwise we would go from November (11) to January (1).

If a calculation uses invalid dates, the calculation fails and produces warnings:

mysql> SELECT '2018-10-31' + INTERVAL 1 DAY;
+-------------------------------+
| '2018-10-31' + INTERVAL 1 DAY |
| 2018-11-01                    |
+-------------------------------+

mysql> SELECT '2018-10-32' + INTERVAL 1 DAY;
+-------------------------------+
| '2018-10-32' + INTERVAL 1 DAY |
| NULL                          |
+-------------------------------+

mysql> SHOW WARNINGS;
+---------+------+----------------------------------------+
| Level   | Code | Message                                |
|---------+------|----------------------------------------+
| Warning | 1292 | Incorrect datetime value: '2018-10-32' |

4.4.6 Working with NULL Values

The NULL value can be surprising until you get used to it. Conceptually, NULL means “a missing unknown value” and it is treated somewhat differently from other values.

To test for NULL, use the IS NULL and IS NOT NULL operators, as shown here:

mysql> SELECT 1 IS NULL, 1 IS NOT NULL;
+-----------+---------------+
| 1 IS NULL | 1 IS NOT NULL |
|         0 |             1 |
+-----------+---------------+

You cannot use arithmetic comparison operators such as =, <, or <> to test for NULL. To demonstrate this for yourself, try the following query:

mysql> SELECT 1 = NULL, 1 <> NULL, 1 < NULL, 1 > NULL;
+----------+-----------+----------+----------+
| 1 = NULL | 1 <> NULL | 1 < NULL | 1 > NULL |
|     NULL |      NULL |     NULL |     NULL |
+----------+-----------+----------+----------+

Because the result of any arithmetic comparison with NULL is also NULL, you cannot obtain any meaningful results from such comparisons.

In MySQL, 0 or NULL means false and anything else means true. The default truth value from a boolean operation is 1.

This special treatment of NULL is why, in the previous section, it was necessary to determine which animals are no longer alive using death IS NOT NULL instead of death <> NULL.

Two NULL values are regarded as equal in a GROUP BY.
Pattern Matching

When doing an ORDER BY, NULL values are presented first if you do ORDER BY ... ASC and last if you do ORDER BY ... DESC.

A common error when working with NULL is to assume that it is not possible to insert a zero or an empty string into a column defined as NOT NULL, but this is not the case. These are in fact values, whereas NULL means “not having a value.” You can test this easily enough by using IS [NOT] NULL as shown:

```sql
mysql> SELECT 0 IS NULL, 0 IS NOT NULL, '' IS NULL, '' IS NOT NULL;
```

Thus it is entirely possible to insert a zero or empty string into a NOT NULL column, as these are in fact NOT NULL. See Problems with NULL Values.

4.4.7 Pattern Matching

MySQL provides standard SQL pattern matching as well as a form of pattern matching based on extended regular expressions similar to those used by Unix utilities such as vi, grep, and sed.

SQL pattern matching enables you to use _ to match any single character and % to match an arbitrary number of characters (including zero characters). In MySQL, SQL patterns are case-insensitive by default. Some examples are shown here. Do not use = or <> when you use SQL patterns. Use the LIKE or NOT LIKE comparison operators instead.

To find names beginning with `b`:

```sql
mysql> SELECT * FROM pet WHERE name LIKE 'b%';
```

| name   | owner  | species | sex | birth      | death      |
|--------|--------|---------|-----|------------+------------|
| Buffy  | Harold | dog     | f   | 1989-05-13 | NULL       |
| Bowser | Diane  | dog     | m   | 1989-08-31 | 1995-07-29 |

To find names ending with `fy`:

```sql
mysql> SELECT * FROM pet WHERE name LIKE '%fy';
```

| name   | owner  | species | sex | birth      | death |
|--------|--------|---------|-----|------------+-------|
| Fluffy | Harold | cat     | f   | 1993-02-04 | NULL  |
| Buffy  | Harold | dog     | f   | 1989-05-13 | NULL  |

To find names containing a `w`:

```sql
mysql> SELECT * FROM pet WHERE name LIKE '%w%';
```

| name   | owner | species | sex | birth      | death          |
|--------|-------|---------|-----|------------+----------------|
| Claws  | Gwen  | cat     | m   | 1994-03-17 | NULL           |
| Bowser | Diane | dog     | m   | 1989-08-31 | 1995-07-29     |
| Whistler | Gwen | bird    | NULL | 1997-12-09 | NULL           |

To find names containing exactly five characters, use five instances of the _ pattern character:

```sql
mysql> SELECT * FROM pet WHERE name LIKE '_____';
```

| name  | owner | species | sex | birth      | death |
|-------|-------|---------|-----|------------+-------|
| Claws | Gwen  | cat     | m   | 1994-03-17 | NULL  |
The other type of pattern matching provided by MySQL uses extended regular expressions. When you test for a match for this type of pattern, use the `REGEXP_LIKE()` function (or the `REGEXP` or `RLIKE` operators, which are synonyms for `REGEXP_LIKE()`).

The following list describes some characteristics of extended regular expressions:

- . matches any single character.
- A character class `[...]` matches any character within the brackets. For example, `[abc]` matches `a`, `b`, or `c`. To name a range of characters, use a dash. `[a-z]` matches any letter, whereas `[0-9]` matches any digit.
- `*` matches zero or more instances of the thing preceding it. For example, `x*` matches any number of `x` characters, `[0-9]*` matches any number of digits, and `. *` matches any number of anything.
- A regular expression pattern match succeeds if the pattern matches anywhere in the value being tested. (This differs from a `LIKE` pattern match, which succeeds only if the pattern matches the entire value.)
- To anchor a pattern so that it must match the beginning or end of the value being tested, use `^` at the beginning or `$` at the end of the pattern.

To demonstrate how extended regular expressions work, the `LIKE` queries shown previously are rewritten here to use `REGEXP_LIKE()`.

To find names beginning with `b`, use `^` to match the beginning of the name:

```sql
mysql> SELECT * FROM pet WHERE REGEXP_LIKE(name, '^b');
+--------+--------+---------+------+------------+------------+
| name   | owner  | species | sex  | birth      | death      |
+--------+--------+---------+------+------------+------------+
| Buffy  | Harold | dog     | f    | 1989-05-13 | NULL       |
| Bowser | Diane  | dog     | m    | 1979-08-31 | 1995-07-29 |
+--------+--------+---------+------+------------+------------+
```

To force a regular expression comparison to be case-sensitive, use a case-sensitive collation, or use the `BINARY` keyword to make one of the strings a binary string, or specify the `c` match-control character. Each of these queries matches only lowercase `b` at the beginning of a name:

```sql
SELECT * FROM pet WHERE REGEXP_LIKE(name, 'b' COLLATE utf8mb4_0900_as_cs);
SELECT * FROM pet WHERE REGEXP_LIKE(name, BINARY 'b');
SELECT * FROM pet WHERE REGEXP_LIKE(name, 'b', 'c');
```

To find names ending with `fy`, use `$` to match the end of the name:

```sql
mysql> SELECT * FROM pet WHERE REGEXP_LIKE(name, 'fy$');
+--------+--------+---------+------+------------+-------+
| name   | owner  | species | sex  | birth      | death |
+--------+--------+---------+------+------------+-------+
| Fluffy | Harold | cat     | f    | 1993-02-04 | NULL  |
| Buffy  | Harold | dog     | f    | 1989-05-13 | NULL  |
+--------+--------+---------+------+------------+-------+
```

To find names containing a `w`, use this query:

```sql
mysql> SELECT * FROM pet WHERE REGEXP_LIKE(name, 'w');
+----------+-------+---------+------+------------+------------+
| name     | owner | species | sex  | birth      | death      |
+----------+-------+---------+------+------------+------------+
| Claws    | Gwen  | cat     | m    | 1994-03-17 | NULL       |
| Bowser   | Diane | dog     | m    | 1989-08-31 | 1995-07-29 |
| Whistler | Gwen  | bird    | NULL | 1997-12-09 | NULL       |
+----------+-------+---------+------+------------+------------+
```
Counting Rows

Because a regular expression pattern matches if it occurs anywhere in the value, it is not necessary in
the previous query to put a wildcard on either side of the pattern to get it to match the entire value as
would be true with an SQL pattern.

To find names containing exactly five characters, use ^ and $ to match the beginning and end of the
name, and five instances of . in between:

```
mysql> SELECT * FROM pet WHERE REGEXP_LIKE(name, '^..\.\.\.$');
+-----+-----+-----+-----+---------+-------+
| name | owner | species | sex | birth   | death |
+-----+-----+-----+-----+---------+-------+
| Claws | Gwen | cat | m   | 1994-03-17 | NULL  |
+-----+-----+-----+-----+---------+-------+
| Buffy | Harold | dog | f   | 1994-03-17 | NULL  |
+-----+
```

You could also write the previous query using the \{n\} ("repeat-n-times") operator:

```
mysql> SELECT * FROM pet WHERE REGEXP_LIKE(name, '^.{5}$');
+-----+-----+-----+-----+---------+-------+
| name | owner | species | sex | birth   | death |
+-----+-----+-----+-----+---------+-------+
| Claws | Gwen | cat | m   | 1994-03-17 | NULL  |
+-----+-----+-----+-----+---------+-------+
| Buffy | Harold | dog | f   | 1994-03-17 | NULL  |
+-----+
```

For more information about the syntax for regular expressions, see Regular Expressions.

### 4.4.8 Counting Rows

Databases are often used to answer the question, “How often does a certain type of data occur in a
table?” For example, you might want to know how many pets you have, or how many pets each owner
has, or you might want to perform various kinds of census operations on your animals.

Counting the total number of animals you have is the same question as “How many rows are in the pet
table?” because there is one record per pet. \( \text{COUNT(*)} \) counts the number of rows, so the query to
count your animals looks like this:

```
mysql> SELECT COUNT(*) FROM pet;
+----------+
| COUNT(*) |
+----------+
|        9 |
+----------+
```

Earlier, you retrieved the names of the people who owned pets. You can use \( \text{COUNT()} \) if you want to
find out how many pets each owner has:

```
mysql> SELECT owner, COUNT(*) FROM pet GROUP BY owner;
+-----+----------+
| owner | COUNT(*) |
+-----+----------+
| Benny |        2 |
| Diane |        2 |
| Gwen  |        3 |
| Harold|        2 |
+-----+
```

The preceding query uses \( \text{GROUP BY} \) to group all records for each \( \text{owner} \). The use of \( \text{COUNT()} \)
in conjunction with \( \text{GROUP BY} \) is useful for characterizing your data under various groupings. The
following examples show different ways to perform animal census operations.

Number of animals per species:

```
mysql> SELECT species, COUNT(*) FROM pet GROUP BY species;
+-------+----------+
| species | COUNT(*) |
+-------+----------+
```

## Counting Rows

| bird    | 2 |
| cat     | 2 |
| dog     | 3 |
| hamster | 1 |
| snake   | 1 |

### Number of animals per sex:

```sql
mysql> SELECT sex, COUNT(*) FROM pet GROUP BY sex;
```

<table>
<thead>
<tr>
<th>sex</th>
<th>COUNT(*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NULL</td>
<td>1</td>
</tr>
<tr>
<td>f</td>
<td>4</td>
</tr>
<tr>
<td>m</td>
<td>4</td>
</tr>
</tbody>
</table>

(In this output, **NULL** indicates that the sex is unknown.)

### Number of animals per combination of species and sex:

```sql
mysql> SELECT species, sex, COUNT(*) FROM pet GROUP BY species, sex;
```

<table>
<thead>
<tr>
<th>species</th>
<th>sex</th>
<th>COUNT(*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bird</td>
<td>NULL</td>
<td>1</td>
</tr>
<tr>
<td>bird</td>
<td>f</td>
<td>1</td>
</tr>
<tr>
<td>cat</td>
<td>f</td>
<td>1</td>
</tr>
<tr>
<td>cat</td>
<td>m</td>
<td>1</td>
</tr>
<tr>
<td>dog</td>
<td>f</td>
<td>1</td>
</tr>
<tr>
<td>dog</td>
<td>m</td>
<td>2</td>
</tr>
<tr>
<td>hamster</td>
<td>f</td>
<td>1</td>
</tr>
<tr>
<td>snake</td>
<td>m</td>
<td>1</td>
</tr>
</tbody>
</table>

You need not retrieve an entire table when you use `COUNT()`. For example, the previous query, when performed just on dogs and cats, looks like this:

```sql
mysql> SELECT species, sex, COUNT(*) FROM pet WHERE species = 'dog' OR species = 'cat' GROUP BY species, sex;
```

<table>
<thead>
<tr>
<th>species</th>
<th>sex</th>
<th>COUNT(*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cat</td>
<td>f</td>
<td>1</td>
</tr>
<tr>
<td>cat</td>
<td>m</td>
<td>1</td>
</tr>
<tr>
<td>dog</td>
<td>f</td>
<td>1</td>
</tr>
<tr>
<td>dog</td>
<td>m</td>
<td>2</td>
</tr>
</tbody>
</table>

Or, if you wanted the number of animals per sex only for animals whose sex is known:

```sql
mysql> SELECT species, sex, COUNT(*) FROM pet WHERE sex IS NOT NULL GROUP BY species, sex;
```

<table>
<thead>
<tr>
<th>species</th>
<th>sex</th>
<th>COUNT(*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bird</td>
<td>f</td>
<td>1</td>
</tr>
<tr>
<td>cat</td>
<td>f</td>
<td>1</td>
</tr>
<tr>
<td>cat</td>
<td>m</td>
<td>1</td>
</tr>
<tr>
<td>dog</td>
<td>f</td>
<td>1</td>
</tr>
<tr>
<td>dog</td>
<td>m</td>
<td>2</td>
</tr>
<tr>
<td>hamster</td>
<td>f</td>
<td>1</td>
</tr>
<tr>
<td>snake</td>
<td>m</td>
<td>1</td>
</tr>
</tbody>
</table>

If you name columns to select in addition to the `COUNT()` value, a `GROUP BY` clause should be present that names those same columns. Otherwise, the following occurs:
• If the `ONLY_FULL_GROUP_BY` SQL mode is enabled, an error occurs:

```
mysql> SET sql_mode = 'ONLY_FULL_GROUP_BY';
Query OK, 0 rows affected (0.00 sec)
mysql> SELECT owner, COUNT(*) FROM pet;
ERROR 1140 (42000): In aggregated query without GROUP BY, expression
#1 of SELECT list contains nonaggregated column 'menagerie.pet.owner';
this is incompatible with sql_mode=only_full_group_by
```

• If `ONLY_FULL_GROUP_BY` is not enabled, the query is processed by treating all rows as a single group, but the value selected for each named column is nondeterministic. The server is free to select the value from any row:

```
mysql> SET sql_mode = '';
Query OK, 0 rows affected (0.00 sec)
mysql> SELECT owner, COUNT(*) FROM pet;
+--------+----------+
<table>
<thead>
<tr>
<th>owner</th>
<th>COUNT(*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harold</td>
<td></td>
</tr>
</tbody>
</table>
+--------+----------+
1 row in set (0.00 sec)
```

See also MySQL Handling of GROUP BY. See Aggregate Function Descriptions for information about `COUNT(expr)` behavior and related optimizations.

### 4.4.9 Using More Than one Table

The `pet` table keeps track of which pets you have. If you want to record other information about them, such as events in their lives like visits to the vet or when litters are born, you need another table. What should this table look like? It needs to contain the following information:

- The pet name so that you know which animal each event pertains to.
- A date so that you know when the event occurred.
- A field to describe the event.
- An event type field, if you want to be able to categorize events.

Given these considerations, the `CREATE TABLE` statement for the `event` table might look like this:

```
mysql> CREATE TABLE event (name VARCHAR(20), date DATE, type VARCHAR(15), remark VARCHAR(255));
```

As with the `pet` table, it is easiest to load the initial records by creating a tab-delimited text file containing the following information:

<table>
<thead>
<tr>
<th>name</th>
<th>date</th>
<th>type</th>
<th>remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluffy</td>
<td>1995-05-15</td>
<td>litter</td>
<td>4 kittens, 3 female, 1 male</td>
</tr>
<tr>
<td>Buffy</td>
<td>1993-06-23</td>
<td>litter</td>
<td>5 puppies, 2 female, 3 male</td>
</tr>
<tr>
<td>Buffy</td>
<td>1994-06-19</td>
<td>litter</td>
<td>3 puppies, 3 female</td>
</tr>
<tr>
<td>Chirpy</td>
<td>1999-03-21</td>
<td>vet</td>
<td>needed beak straightened</td>
</tr>
<tr>
<td>Slim</td>
<td>1997-08-03</td>
<td>vet</td>
<td>broken rib</td>
</tr>
<tr>
<td>Bowser</td>
<td>1991-10-12</td>
<td>kennel</td>
<td></td>
</tr>
<tr>
<td>Fang</td>
<td>1991-10-12</td>
<td>kennel</td>
<td></td>
</tr>
<tr>
<td>Fang</td>
<td>1998-08-28</td>
<td>birthday</td>
<td>Gave him a new chew toy</td>
</tr>
</tbody>
</table>
Using More Than One Table

<table>
<thead>
<tr>
<th>name</th>
<th>date</th>
<th>type</th>
<th>remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Claws</td>
<td>1998-03-17</td>
<td>birthday</td>
<td>Gave him a new flea collar</td>
</tr>
<tr>
<td>Whistler</td>
<td>1998-12-09</td>
<td>birthday</td>
<td>First birthday</td>
</tr>
</tbody>
</table>

Load the records like this:

```sql
mysql> LOAD DATA LOCAL INFILE 'event.txt' INTO TABLE event;
```

Based on what you have learned from the queries that you have run on the `pet` table, you should be able to perform retrievals on the records in the `event` table; the principles are the same. But when is the `event` table by itself insufficient to answer questions you might ask?

Suppose that you want to find out the ages at which each pet had its litters. We saw earlier how to calculate ages from two dates. The litter date of the mother is in the `event` table, but to calculate her age on that date you need her birth date, which is stored in the `pet` table. This means the query requires both tables:

```sql
mysql> SELECT pet.name, 
          TIMESTAMPDIFF(YEAR,birth,date) AS age, 
          remark 
FROM pet INNER JOIN event 
ON pet.name = event.name 
WHERE event.type = 'litter';
```

<table>
<thead>
<tr>
<th>name</th>
<th>age</th>
<th>remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluffy</td>
<td>2</td>
<td>4 kittens, 3 female, 1 male</td>
</tr>
<tr>
<td>Buffy</td>
<td>4</td>
<td>5 puppies, 2 female, 3 male</td>
</tr>
<tr>
<td>Buffy</td>
<td>5</td>
<td>3 puppies, 3 female</td>
</tr>
</tbody>
</table>

There are several things to note about this query:

- The **FROM** clause joins two tables because the query needs to pull information from both of them.

- When combining (joining) information from multiple tables, you need to specify how records in one table can be matched to records in the other. This is easy because they both have a **name** column. The query uses an **ON** clause to match up records in the two tables based on the **name** values.

  The query uses an **INNER JOIN** to combine the tables. An **INNER JOIN** permits rows from either table to appear in the result if and only if both tables meet the conditions specified in the **ON** clause. In this example, the **ON** clause specifies that the **name** column in the `pet` table must match the **name** column in the `event` table. If a name appears in one table but not the other, the row does not appear in the result because the condition in the **ON** clause fails.

- Because the **name** column occurs in both tables, you must be specific about which table you mean when referring to the column. This is done by prepending the table name to the column name.

You need not have two different tables to perform a join. Sometimes it is useful to join a table to itself, if you want to compare records in a table to other records in that same table. For example, to find breeding pairs among your pets, you can join the `pet` table with itself to produce candidate pairs of live males and females of like species:

```sql
mysql> SELECT p1.name, p1.sex, p2.name, p2.sex, p1.species 
FROM pet AS p1 INNER JOIN pet AS p2 
ON p1.species = p2.species 
AND p1.sex = 'f' AND p1.death IS NULL 
AND p2.sex = 'm' AND p2.death IS NULL;
```

<table>
<thead>
<tr>
<th>name</th>
<th>sex</th>
<th>name</th>
<th>sex</th>
<th>species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluffy</td>
<td>f</td>
<td>Claws</td>
<td>m</td>
<td>cat</td>
</tr>
<tr>
<td>Buffy</td>
<td>f</td>
<td>Fang</td>
<td>m</td>
<td>dog</td>
</tr>
</tbody>
</table>
In this query, we specify aliases for the table name to refer to the columns and keep straight which instance of the table each column reference is associated with.
Chapter 5 Getting Information About Databases and Tables

What if you forget the name of a database or table, or what the structure of a given table is (for example, what its columns are called)? MySQL addresses this problem through several statements that provide information about the databases and tables it supports.

You have previously seen `SHOW DATABASES`, which lists the databases managed by the server. To find out which database is currently selected, use the `DATABASE()` function:

```
mysql> SELECT DATABASE();
+------------+
| DATABASE() |
+------------+
| menagerie  |
+------------+
```

If you have not yet selected any database, the result is `NULL`.

To find out what tables the default database contains (for example, when you are not sure about the name of a table), use this statement:

```
mysql> SHOW TABLES;
+---------------------+
| Tables_in_menagerie |
+---------------------+
| event               |
| pet                 |
+---------------------+
```

The name of the column in the output produced by this statement is always `Tables_in_db_name`, where `db_name` is the name of the database. See `SHOW TABLES Statement`, for more information.

If you want to find out about the structure of a table, the `DESCRIBE` statement is useful; it displays information about each of a table's columns:

```
mysql> DESCRIBE pet;
+---------+-------------+------+-----+---------+-------+
| Field   | Type        | Null | Key | Default | Extra |
+---------+-------------+------+-----+---------+-------+
| name    | varchar(20) | YES  |     | NULL    |       |
| owner   | varchar(20) | YES  |     | NULL    |       |
| species | varchar(20) | YES  |     | NULL    |       |
| sex     | char(1)     | YES  |     | NULL    |       |
| birth   | date        | YES  |     | NULL    |       |
| death   | date        | YES  |     | NULL    |       |
+---------+-------------+------+-----+---------+-------+
```

Field indicates the column name, Type is the data type for the column, NULL indicates whether the column can contain NULL values, Key indicates whether the column is indexed, and Default specifies the column's default value. Extra displays special information about columns: If a column was created with the `AUTO_INCREMENT` option, the value is `auto_increment` rather than empty.

`DESC` is a short form of `DESCRIBE`. See `DESCRIBE Statement`, for more information.

You can obtain the `CREATE TABLE` statement necessary to create an existing table using the `SHOW CREATE TABLE` statement. See `SHOW CREATE TABLE Statement`.

If you have indexes on a table, `SHOW INDEX FROM tbl_name` produces information about them. See `SHOW INDEX Statement`, for more about this statement.
Chapter 6 Using mysql in Batch Mode

In the previous sections, you used mysql interactively to enter statements and view the results. You can also run mysql in batch mode. To do this, put the statements you want to run in a file, then tell mysql to read its input from the file:

```bash
$> mysql < batch-file
```

If you are running mysql under Windows and have some special characters in the file that cause problems, you can do this:

```bash
C:\> mysql -e "source batch-file"
```

If you need to specify connection parameters on the command line, the command might look like this:

```bash
$> mysql -h host -u user -p < batch-file
Enter password: ********
```

When you use mysql this way, you are creating a script file, then executing the script.

If you want the script to continue even if some of the statements in it produce errors, you should use the --force command-line option.

Why use a script? Here are a few reasons:

- If you run a query repeatedly (say, every day or every week), making it a script enables you to avoid retyping it each time you execute it.
- You can generate new queries from existing ones that are similar by copying and editing script files.
- Batch mode can also be useful while you're developing a query, particularly for multiple-line statements or multiple-statement sequences. If you make a mistake, you don't have to retype everything. Just edit your script to correct the error, then tell mysql to execute it again.
- If you have a query that produces a lot of output, you can run the output through a pager rather than watching it scroll off the top of your screen:

```bash
$> mysql < batch-file | more
```

- You can catch the output in a file for further processing:

```bash
$> mysql < batch-file > mysql.out
```

- You can distribute your script to other people so that they can also run the statements.
- Some situations do not allow for interactive use, for example, when you run a query from a cron job. In this case, you must use batch mode.

The default output format is different (more concise) when you run mysql in batch mode than when you use it interactively. For example, the output of `SELECT DISTINCT species FROM pet` looks like this when mysql is run interactively:

```plaintext
+---------+
| species |
+---------+
| bird    |
| cat     |
| dog     |
| hamster |
| snake   |
+---------+
```

In batch mode, the output looks like this instead:

```plaintext
species
```
If you want to get the interactive output format in batch mode, use `mysql -t`. To echo to the output the statements that are executed, use `mysql -v`.

You can also use scripts from the `mysql` prompt by using the `source` command or `. filename` command:

```sql
mysql> source filename;
mysql> . filename
```

See Executing SQL Statements from a Text File, for more information.
Here are examples of how to solve some common problems with MySQL.

Some of the examples use the table shop to hold the price of each article (item number) for certain traders (dealers). Supposing that each trader has a single fixed price per article, then (article, dealer) is a primary key for the records.

Start the command-line tool `mysql` and select a database:

```
\$> mysql your-database-name
```

To create and populate the example table, use these statements:

```sql
CREATE TABLE shop (
    article INT UNSIGNED  DEFAULT '0000' NOT NULL,
    dealer  CHAR(20)      DEFAULT ''     NOT NULL,
    price   DECIMAL(16,2) DEFAULT '0.00' NOT NULL,
    PRIMARY KEY(article, dealer));
INSERT INTO shop VALUES
    (1,'A',3.45),(1,'B',3.99),(2,'A',10.99),(3,'B',1.45),
    (3,'C',1.69),(3,'D',1.25),(4,'D',19.95);
```

After issuing the statements, the table should have the following contents:

```
<table>
<thead>
<tr>
<th>article</th>
<th>dealer</th>
<th>price</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>3.45</td>
</tr>
<tr>
<td>1</td>
<td>B</td>
<td>3.99</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>10.99</td>
</tr>
<tr>
<td>3</td>
<td>B</td>
<td>1.45</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>1.69</td>
</tr>
<tr>
<td>3</td>
<td>D</td>
<td>1.25</td>
</tr>
<tr>
<td>4</td>
<td>D</td>
<td>19.95</td>
</tr>
</tbody>
</table>
```

### 7.1 The Maximum Value for a Column

“What is the highest item number?”

```sql
SELECT MAX(article) AS article FROM shop;
```

```
<table>
<thead>
<tr>
<th>article</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
</tr>
</tbody>
</table>
```
7.2 The Row Holding the Maximum of a Certain Column

Task: Find the number, dealer, and price of the most expensive article.

This is easily done with a subquery:

```
SELECT article, dealer, price
FROM   shop
WHERE  price=(SELECT MAX(price) FROM shop);
```

<table>
<thead>
<tr>
<th>article</th>
<th>dealer</th>
<th>price</th>
</tr>
</thead>
<tbody>
<tr>
<td>0004</td>
<td>D</td>
<td>19.95</td>
</tr>
</tbody>
</table>

Another solution is to use a `LEFT JOIN`, as shown here:

```
SELECT s1.article, s1.dealer, s1.price
FROM shop s1
LEFT JOIN shop s2 ON s1.price < s2.price
WHERE s2.article IS NULL;
```

You can also do this by sorting all rows descending by price and get only the first row using the MySQL-specific `LIMIT` clause, like this:

```
SELECT article, dealer, price
FROM shop
ORDER BY price DESC
LIMIT 1;
```

**Note**

If there were several most expensive articles, each with a price of 19.95, the `LIMIT` solution would show only one of them.

7.3 Maximum of Column per Group

Task: Find the highest price per article.

```
SELECT article, MAX(price) AS price
FROM   shop
GROUP BY article
ORDER BY article;
```

<table>
<thead>
<tr>
<th>article</th>
<th>price</th>
</tr>
</thead>
<tbody>
<tr>
<td>0001</td>
<td>3.99</td>
</tr>
<tr>
<td>0002</td>
<td>10.99</td>
</tr>
<tr>
<td>0003</td>
<td>1.69</td>
</tr>
<tr>
<td>0004</td>
<td>19.95</td>
</tr>
</tbody>
</table>

7.4 The Rows Holding the Group-wise Maximum of a Certain Column

Task: For each article, find the dealer or dealers with the most expensive price.

This problem can be solved with a subquery like this one:

```
SELECT article, dealer, price
FROM   shop s1
WHERE  price=(SELECT MAX(s2.price)
             FROM shop s2
             WHERE s1.article = s2.article)
ORDER BY article;
```

<table>
<thead>
<tr>
<th>article</th>
<th>dealer</th>
<th>price</th>
</tr>
</thead>
</table>
The preceding example uses a correlated subquery, which can be inefficient (see Correlated Subqueries). Other possibilities for solving the problem are to use an uncorrelated subquery in the FROM clause, a LEFT JOIN, or a common table expression with a window function.

Uncorrelated subquery:

```sql
SELECT s1.article, dealer, s1.price
FROM shop s1
JOIN (SELECT article, MAX(price) AS price
      FROM shop
      GROUP BY article) AS s2
ON s1.article = s2.article AND s1.price = s2.price
ORDER BY article;
```

LEFT JOIN:

```sql
SELECT s1.article, s1.dealer, s1.price
FROM shop s1
LEFT JOIN shop s2 ON s1.article = s2.article AND s1.price < s2.price
WHERE s2.article IS NULL
ORDER BY s1.article;
```

The LEFT JOIN works on the basis that when s1.price is at its maximum value, there is no s2.price with a greater value and thus the corresponding s2.article value is NULL. See JOIN Clause.

Common table expression with window function:

```sql
WITH s1 AS (SELECT article, dealer, price,
             RANK() OVER (PARTITION BY article
                          ORDER BY price DESC
                      ) AS `Rank`
             FROM shop
         )
SELECT article, dealer, price
FROM s1
WHERE `Rank` = 1
ORDER BY article;
```

### 7.5 Using User-Defined Variables

You can employ MySQL user variables to remember results without having to store them in temporary variables in the client. (See User-Defined Variables.)

For example, to find the articles with the highest and lowest price you can do this:

```sql
mysql> SELECT @min_price:=MIN(price),@max_price:=MAX(price) FROM shop;
mysql> SELECT * FROM shop WHERE price=@min_price OR price=@max_price;
```

<table>
<thead>
<tr>
<th>article</th>
<th>dealer</th>
<th>price</th>
</tr>
</thead>
<tbody>
<tr>
<td>0003</td>
<td>D</td>
<td>1.25</td>
</tr>
<tr>
<td>0004</td>
<td>D</td>
<td>19.95</td>
</tr>
</tbody>
</table>

**Note**

It is also possible to store the name of a database object such as a table or a column in a user variable and then to use this variable in an SQL statement;
Using Foreign Keys

MySQL supports foreign keys, which permit cross-referencing related data across tables, and foreign key constraints, which help keep the related data consistent.

A foreign key relationship involves a parent table that holds the initial column values, and a child table with column values that reference the parent column values. A foreign key constraint is defined on the child table.

This following example relates parent and child tables through a single-column foreign key and shows how a foreign key constraint enforces referential integrity.

Create the parent and child tables:

```sql
CREATE TABLE parent (
    id INT NOT NULL,
    PRIMARY KEY (id)
) ENGINE=INNODB;
CREATE TABLE child (
    id INT,
    parent_id INT,
    INDEX par_ind (parent_id),
    FOREIGN KEY (parent_id)
    REFERENCES parent(id)
) ENGINE=INNODB;
```

Insert a row into the parent table:

```sql
mysql> INSERT INTO parent (id) VALUES (1);
```

Verify that the data was inserted:

```sql
mysql> SELECT * FROM parent;
+----+
| id |
+----+
|  1 |
```

Insert a row into the child table:

```sql
mysql> INSERT INTO child (id,parent_id) VALUES (1,1);
```

The insert operation is successful because parent_id 1 is present in the parent table.

Insert a row into the child table with a parent_id value that is not present in the parent table:

```sql
mysql> INSERT INTO child (id,parent_id) VALUES(2,2);
ERROR 1452 (23000): Cannot add or update a child row: a foreign key constraint fails
('test'.child', CONSTRAINT 'child_ibfk_1' FOREIGN KEY ('parent_id')
REFERENCES 'parent' ('id'))
```

The operation fails because the specified parent_id value does not exist in the parent table.

Try to delete the previously inserted row from the parent table:

```sql
mysql> DELETE FROM parent WHERE id VALUES = 1;
ERROR 1451 (23000): Cannot delete or update a parent row: a foreign key constraint fails
('test'.child', CONSTRAINT 'child_ibfk_1' FOREIGN KEY ('parent_id')
REFERENCES 'parent' ('id'))
```

This operation fails because the record in the child table contains the referenced id (parent_id) value.
When an operation affects a key value in the parent table that has matching rows in the child table, the result depends on the referential action specified by \texttt{ON UPDATE} and \texttt{ON DELETE} subclauses of the \texttt{FOREIGN KEY} clause. Omitting \texttt{ON DELETE} and \texttt{ON UPDATE} clauses (as in the current child table definition) is the same as specifying the \texttt{RESTRICT} option, which rejects operations that affect a key value in the parent table that has matching rows in the parent table.

To demonstrate \texttt{ON DELETE} and \texttt{ON UPDATE} referential actions, drop the child table and recreate it to include \texttt{ON UPDATE} and \texttt{ON DELETE} subclauses with the \texttt{CASCADE} option. The \texttt{CASCADE} option automatically deletes or updates matching rows in the child table when deleting or updating rows in the parent table.

```sql
DROP TABLE child;
CREATE TABLE child (id INT, parent_id INT, INDEX par_ind (parent_id), FOREIGN KEY (parent_id) REFERENCES parent(id) ON UPDATE CASCADE ON DELETE CASCADE) ENGINE=INNODB;
```

Insert the following rows into the child table:

```sql
mysql> INSERT INTO child (id,parent_id) VALUES(1,1),(2,1),(3,1);
```

Verify that the data was inserted:

```sql
mysql> SELECT * FROM child;
```

```
+------+-----------+
<table>
<thead>
<tr>
<th>id</th>
<th>parent_id</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>
+------+-----------+
```

Update the id in the parent table, changing it from 1 to 2.

```sql
mysql> UPDATE parent SET id = 2 WHERE id = 1;
```

Verify that the update was successful:

```sql
mysql> SELECT * FROM parent;
```

```
+----+
| id |
+----+
| 2  |
+----+
```

Verify that the \texttt{ON UPDATE CASCADE} referential action updated the child table:

```sql
mysql> SELECT * FROM child;
```

```
+------+-----------+
<table>
<thead>
<tr>
<th>id</th>
<th>parent_id</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>
+------+-----------+
```

To demonstrate the \texttt{ON DELETE CASCADE} referential action, delete records from the parent table where the \texttt{parent_id} = 2, which deletes all records in the parent table.

```sql
mysql> DELETE FROM parent WHERE id = 2;
```

Because all records in the child table are associated with \texttt{parent_id} = 2, the \texttt{ON DELETE CASCADE} referential action removes all records from the child table.
For more information about foreign key constraints, see FOREIGN KEY Constraints.

7.7 Searching on Two Keys

An OR using a single key is well optimized, as is the handling of AND.

The one tricky case is that of searching on two different keys combined with OR:

```
SELECT field1_index, field2_index FROM test_table
WHERE field1_index = '1' OR field2_index = '1'
```

This case is optimized. See Index Merge Optimization.

You can also solve the problem efficiently by using a UNION that combines the output of two separate SELECT statements. See UNION Clause.

Each SELECT searches only one key and can be optimized:

```
SELECT field1_index, field2_index
FROM test_table WHERE field1_index = '1'
UNION
SELECT field1_index, field2_index
FROM test_table WHERE field2_index = '1';
```

7.8 Calculating Visits Per Day

The following example shows how you can use the bit group functions to calculate the number of days per month a user has visited a Web page.

```
CREATE TABLE t1 (year YEAR, month INT UNSIGNED, day INT UNSIGNED);
INSERT INTO t1 VALUES(2000,1,1),(2000,1,20),(2000,1,30),(2000,2,2),
(2000,2,23),(2000,2,23);
```

The example table contains year-month-day values representing visits by users to the page. To determine how many different days in each month these visits occur, use this query:

```
SELECT year,month,BIT_COUNT(BIT_OR(1<<day)) AS days FROM t1
GROUP BY year,month;
```

Which returns:

```
+------+-------+------+
<table>
<thead>
<tr>
<th>year</th>
<th>month</th>
<th>days</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2000</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
+------+-------+------+
```

The query calculates how many different days appear in the table for each year/month combination, with automatic removal of duplicate entries.

7.9 Using AUTO_INCREMENT

The AUTO_INCREMENT attribute can be used to generate a unique identity for new rows:

```
CREATE TABLE animals {
    id MEDIUMINT NOT NULL AUTO_INCREMENT,
    name CHAR(30) NOT NULL,
    PRIMARY KEY (id)
};
INSERT INTO animals (name) VALUES
    ('dog'), ('cat'), ('penguin'),
```
Using AUTO_INCREMENT

```
('lax'),('whale'),('ostrich');
SELECT * FROM animals;
```

Which returns:

<table>
<thead>
<tr>
<th>id</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>dog</td>
</tr>
<tr>
<td>2</td>
<td>cat</td>
</tr>
<tr>
<td>3</td>
<td>penguin</td>
</tr>
<tr>
<td>4</td>
<td>lax</td>
</tr>
<tr>
<td>5</td>
<td>whale</td>
</tr>
<tr>
<td>6</td>
<td>ostrich</td>
</tr>
</tbody>
</table>

No value was specified for the `AUTO_INCREMENT` column, so MySQL assigned sequence numbers automatically. You can also explicitly assign 0 to the column to generate sequence numbers, unless the `NO_AUTO_VALUE_ON_ZERO` SQL mode is enabled. For example:

```
INSERT INTO animals (id,name) VALUES(0,'groundhog');
```

If the column is declared `NOT NULL`, it is also possible to assign `NULL` to the column to generate sequence numbers. For example:

```
INSERT INTO animals (id,name) VALUES(NULL,'squirrel');
```

When you insert any other value into an `AUTO_INCREMENT` column, the column is set to that value and the sequence is reset so that the next automatically generated value follows sequentially from the largest column value. For example:

```
INSERT INTO animals (id,name) VALUES(100,'rabbit');
INSERT INTO animals (id,name) VALUES(NULL,'mouse');
SELECT * FROM animals;
```

<table>
<thead>
<tr>
<th>id</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>dog</td>
</tr>
<tr>
<td>2</td>
<td>cat</td>
</tr>
<tr>
<td>3</td>
<td>penguin</td>
</tr>
<tr>
<td>4</td>
<td>lax</td>
</tr>
<tr>
<td>5</td>
<td>whale</td>
</tr>
<tr>
<td>6</td>
<td>ostrich</td>
</tr>
<tr>
<td>7</td>
<td>groundhog</td>
</tr>
<tr>
<td>8</td>
<td>squirrel</td>
</tr>
<tr>
<td>100</td>
<td>rabbit</td>
</tr>
<tr>
<td>101</td>
<td>mouse</td>
</tr>
</tbody>
</table>

Updating an existing `AUTO_INCREMENT` column value also resets the `AUTO_INCREMENT` sequence.

You can retrieve the most recent automatically generated `AUTO_INCREMENT` value with the `LAST_INSERT_ID()` SQL function or the `mysql_insert_id()` C API function. These functions are connection-specific, so their return values are not affected by another connection which is also performing inserts.

Use the smallest integer data type for the `AUTO_INCREMENT` column that is large enough to hold the maximum sequence value you require. When the column reaches the upper limit of the data type, the next attempt to generate a sequence number fails. Use the `UNSIGNED` attribute if possible to allow a greater range. For example, if you use `TINYINT`, the maximum permissible sequence number is 127. For `TINYINT UNSIGNED`, the maximum is 255. See Integer Types (Exact Value) - `INTEGER`, `INT`, `SMALLINT`, `TINYINT`, `MEDIUMINT`, `BIGINT` for the ranges of all the integer types.

**Note**

For a multiple-row insert, `LAST_INSERT_ID()` and `mysql_insert_id()` actually return the `AUTO_INCREMENT` key from the *first* of the inserted rows.
InnoDB Notes

This enables multiple-row inserts to be reproduced correctly on other servers in a replication setup.

To start with an AUTO_INCREMENT value other than 1, set that value with CREATE TABLE or ALTER TABLE, like this:

\`
mysql> ALTER TABLE tbl AUTO_INCREMENT = 100;
\`

InnoDB Notes

For information about AUTO_INCREMENT usage specific to InnoDB, see AUTO_INCREMENT Handling in InnoDB.

MyISAM Notes

- For MyISAM tables, you can specify AUTO_INCREMENT on a secondary column in a multiple-column index. In this case, the generated value for the AUTO_INCREMENT column is calculated as \( \text{MAX(auto_increment_column)} + 1 \) WHERE prefix=given-prefix. This is useful when you want to put data into ordered groups.

```sql
CREATE TABLE animals (  
grp ENUM('fish','mammal','bird') NOT NULL,  
id MEDIUMINT NOT NULL AUTO_INCREMENT,  
name CHAR(30) NOT NULL,  
PRIMARY KEY (grp,id)  
) ENGINE=MyISAM;
```

```sql
INSERT INTO animals (grp,name) VALUES  
('mammal','dog'),('mammal','cat'),  
('bird','penguin'),('fish','lax'),('mammal','whale'),  
('bird','ostrich');
```

```sql
SELECT * FROM animals ORDER BY grp,id;
```

Which returns:

<table>
<thead>
<tr>
<th>grp</th>
<th>id</th>
<th>name</th>
</tr>
</thead>
<tbody>
<tr>
<td>fish</td>
<td>1</td>
<td>lax</td>
</tr>
<tr>
<td>mammal</td>
<td>1</td>
<td>dog</td>
</tr>
<tr>
<td>mammal</td>
<td>2</td>
<td>cat</td>
</tr>
<tr>
<td>mammal</td>
<td>3</td>
<td>whale</td>
</tr>
<tr>
<td>bird</td>
<td>1</td>
<td>penguin</td>
</tr>
<tr>
<td>bird</td>
<td>2</td>
<td>ostrich</td>
</tr>
</tbody>
</table>

In this case (when the AUTO_INCREMENT column is part of a multiple-column index), AUTO_INCREMENT values are reused if you delete the row with the biggest AUTO_INCREMENT value in any group. This happens even for MyISAM tables, for which AUTO_INCREMENT values normally are not reused.

- If the AUTO_INCREMENT column is part of multiple indexes, MySQL generates sequence values using the index that begins with the AUTO_INCREMENT column, if there is one. For example, if the animals table contained indexes PRIMARY KEY (grp, id) and INDEX (id), MySQL would ignore the PRIMARY KEY for generating sequence values. As a result, the table would contain a single sequence, not a sequence per grp value.

Further Reading

More information about AUTO_INCREMENT is available here:

- How to assign the AUTO_INCREMENT attribute to a column: CREATE TABLE Statement, and ALTER TABLE Statement.
- How AUTO_INCREMENT behaves depending on the NO_AUTO_VALUE_ON_ZERO SQL mode: Server SQL Modes.
• How to use the `LAST_INSERT_ID()` function to find the row that contains the most recent `AUTO_INCREMENT` value: Information Functions.

• Setting the `AUTO_INCREMENT` value to be used: Server System Variables.

• `AUTO_INCREMENT` Handling in InnoDB

• `AUTO_INCREMENT` and replication: Replication and AUTO_INCREMENT.

• Server-system variables related to `AUTO_INCREMENT` (auto_increment_increment and auto_increment_offset) that can be used for replication: Server System Variables.
Chapter 8 Using MySQL with Apache

There are programs that let you authenticate your users from a MySQL database and also let you write your log files into a MySQL table.

You can change the Apache logging format to be easily readable by MySQL by putting the following into the Apache configuration file:

```
LogFormat ""%h",%{%Y%m%d%H%M%S}t,.%>s,""%b"",""%{Content-Type}o"", 
""%U"",""%{Referer}i",""%{User-Agent}i"
```

To load a log file in that format into MySQL, you can use a statement something like this:

```
LOAD DATA INFILE '/local/access_log' INTO TABLE tbl_name
FIELDS TERMINATED BY ',' OPTIONALLY ENCLOSED BY '"' ESCAPED BY '\\'
```

The named table should be created to have columns that correspond to those that the LogFormat line writes to the log file.