

Manipulating data.: I have problems with importing CSV data

Problems with importing CSV-files frequently occur at the Webyog Forums. We will here outline the basics.

1) What is CSV?

CSV stands for 'Comma Separated Values'. Although CSV represents some attempt to standardize exchange of data across applications using text files, it is poorly standardized. Sometimes a singlequote is used for enclosing values, sometimes a doublequote; sometimes all types of data are enclosed, sometimes only string data. And in most of continental Europe (where the comma is used for decimal notation - ie: "4.123,7" instead of English notation "4,123.7") it is a de facto convention to use semicolon and not comma as field separator! Localized versions of Microsoft Excel and the Microsoft text-ODBC-driver handles it that way for instance! Thus as CSV really becomes a SSV ('Semicolon Separated Value') in most of Europe. But not matter what: SQLyog handles it all.

2) How does SQLyog import CSV to a MySQL database?

SQLyog issues the MySQL-command '**LOAD DATA LOCAL INFILE**'. Read the MySQL docs on that here: <http://dev.mysql.com/doc/refman/5.0/en/load-data.html>

The values entered in the settings dialogue simply are used for constructing the LOAD DATA LOCAL INFILE statement.

As a consequence you cannot import a CSV-file at your local to a remote MySQL server when connected with HTTP-tunneling - because technically the client is the PHP binary running at the remote host with HTTP-tunneling. So the system would look for the file at the remote host itself! Also MySQL cannot have LOAD DATA LOCAL INFILE disabled with "local-infile = 0" directive in the configuration file (my.cnf/my.ini)

The MySQL-command 'LOAD DATA LOCAL INFILE' is actually the 'inverse' of what SQLyog does when it generates a CSV-file - it then simply executes

"SELECT .. into outfile ... (optionally) enclosed by ... terminated by .. escaped by ..."

3) A note on escaping etc.

1) Starting with version 5.13 a small improvement with the 'Export as CSV' and 'Import as CSV' settings dialogue and the 'standard' MySQL **escape** character '\ ' has been implemented: No matter if you enter '\ ' or '\\ ' as escape character, SQLyog now does the same: it simply uses '\ ' as the escape character. What is sent to MySQL is (literally) " ... escaped by '\\ ' - and as '\\ ' is an escaped '\ ' it will use '\ ' as the escape character (this escaping is exactly as you will have to send "INSERT INTO ... (... ,'backs\ash', ...) " to the server to insert the literal string 'backs\ash'. It has confused quite a lot of users over the years that the escape character '\ ' itself also had to be escaped like '\\ ' in the 'Export as CSV' settings dialogue. As of SQLyog 5.13, you may escape it or not - or put another way: both '\ ' and '\\ ' in the settings dialogue makes SQLyog

Manipulating data.: I have problems with importing CSV data

use the escape character '\'. It sends " ... escaped by \" to MySQL. To allow for the use of \" the settings dialogue lets you input two characters, but you should only use one: the 'standard' \' (entered as \' or \\\) or any other character.

2) you must use the platform-independent C-syntax notation (that also MySQL uses) for "linefeed" (\n), "carriage return" (\r) and "tab" (\t); unless you have changed the escape character - you will then of course need to use your chose n escape character.

3) When using the singlequote character as delimiter when importing strings that have the singlequote character in it itself the **enclose** character must be specified as " \' "

Read the MySQL docs on how to use escaping within strings:

<http://dev.mysql.com/doc/refman/5.0/en/string-syntax.html>

But as said - with the **LOAD DATA INFILE** and **INSERT INTO OUTFILE** statements as generated by SQLyog GUI you may override the standard values (if the word 'standard' applies to CSV at all!) - and it then only has effect with that single statement only. For instance you can use the | character ("pipe") as delimiter, the * character as enclose character and the ^ character as escape character if you want.

Overriding the standards can be practical (makes the file more human readable) when working with character data having comma, semicolon, singlequote and doublequote characters (and maybe even " \' " and more 'problematic' characters!) in them. Provided that the application used 'at the other end' has that flexibility.

4) A few tricks

1) To import a tab-delimited txt-file simply use \t as the delimiter setting. That is a workaround to use with Excel-data where you do not know what characters occurs inside strings and have difficulty in choosing an enclose character: export as a tab-delimited .txt file, and use the setting " \t " as delimiter when importing with SQLyog. It should be clear now that you can import almost any format text file with SQLyog once you understand to 'tune' the settings.

2) Consider using ODBC instead. For instance an Excel data file (.xls -file) can be opened from Microsoft Access. After saving data in Access-format you can transfer data using ODBC. Then you even don't have to define the columns in advance if the spreadsheet contains the column names in the upper row. ODBC does that for you. And with the SQLyog Import External Data Tool you can adjust the datatypes to your need in the import process.

3) Don't miss the settings available at the bottom of the settings dialogue! With the options available here you can skip header - and other garbage - lines, and you can control if existing data shall be overwritten or not. Refer to the SQLyog help file on that!

Also read this FAQ on importing CSV-data using the Import External Data Tool

http://www.webyog.com/faq/14_88_en.html

Manipulating data.: I have problems with importing CSV data

Manipulating data.: I have problems with importing CSV data

Unique solution ID: #1058

Author: Peter Laursen

Last update: 2006-05-30 11:52