

# ISSS615 Data Management

## MySQL LoadData Handout

### Table of Contents

A. Loading data into MySQL Table from Tab-delimited TXT file (or , delimited csv file) via Load Infile command.....	2
B. Frequently asked questions.....	3

## A. Loading data into MySQL Table from Tab-delimited TXT file (or, delimited csv file) via **Load Infile** command

1. Use '**LOAD INFILE**' command to load the data directly from the .txt file.

Detailed steps are listed as follows.

Step 1: Place the .txt file (containing data separated by tabs and carriage return character) in a folder say C:/<folder>/GXTY/Data (e.g., C:/<folder>/GXTY/Data/order.txt).

(you may be placing txt files in a folder allowed by *secur\_file\_priv* MySQL setting)

Step 2: Every .txt file is corresponding to ONLY one table. Like Order.txt may correspond to table order\_t

Step 3: Start WAMP and MYSQL Workbench. Connect to the database into which the tables need to be created.

Step 4: Open a new Query Tab

Step 5: Enter the command: **LOAD DATA INFILE** 'C:/<folder>/GXTY/Data/order.txt' **INTO TABLE** order\_t **FIELDS TERMINATED BY** '\t' **LINES TERMINATED BY** '\r\n' **IGNORE 1 LINES**;

Clauses used with LOAD DATA command:

- Using **FIELDS TERMINATED BY** '\t' identifies tabs as the delimiter of columns and is used to input data into corresponding columns in the database
- Using **FIELDS TERMINATED BY** ',' identifies ,(comma) as the delimiter of columns and is used to input data into corresponding columns in the database
- If you have generated the text file on a Windows system, you might have to use **LINES TERMINATED BY** '\r\n' to read the file properly, because Windows programs typically use these two characters as a line terminator. Some editors like WordPad may just use \r as line terminator. This typically happens when you are on systems that run on a different OS. On Notepad++ - when you enable View-> Show Symbols -> Show all Characters, you will see CRLF('\r\n') or LF('\n') or CR('\r') added to the end of the line. Accordingly, use **LINES TERMINATED BY** '\r\n' or **LINES TERMINATED BY** '\n' or **LINES TERMINATED BY** '\r' as is required.
- **IGNORE 1 LINES** is to ignore the first line in order.txt file that is used to record column names. If your \*.txt file does not contain column names in the first line, the command script should NOT contain '**IGNORE 1 LINES**'.

## B. Frequently asked questions

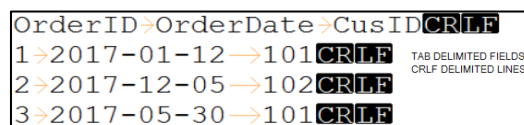
1. I get the error number 1452 with description 'Cannot add or update a child row: a foreign key constraint fails..'

### Answer:

This error occurs when you are inserting data into a table that has Foreign Key (FK) columns and the value from the .txt file that you are inserting is not present in the corresponding Primary Key (PK) table's column that this FK column is referring to.

Some suggestions to help solve this issue:

1. Check the order of the columns in the MySQL table and corresponding .txt file, they should match.
2. Check in the .txt file containing data if you have any empty rows in the end or empty columns in the end of any row (creates problem because an empty value cannot find a match in the PK column)
3. Check if the foreign key column is left empty in 1 or more rows.  
Leaving FK column empty does not indicate to MySQL that **null** has to be inserted.  
Read FAQ #2 to understand how to load 'null' values.
4. Check the data (.txt) file to see what delimiters are used. Use appropriate delimiter clauses "FIELDS TERMINATED BY" and "LINES TERMINATED BY "
5. If the column corresponding to the foreign key is the last column in the .txt file, you may get this error even though the value corresponding the last column exists in the corresponding primary table. Use LINES TERMINATED BY '\r\n' or '\n' or '\r' in LOAD infile command.



```
OrderID>OrderDate>CusIDCRLF
1>2017-01-12->101CRLF
2>2017-12-05->102CRLF
3>2017-05-30->101CRLF
```

TAB DELIMITED FIELDS  
CRLF DELIMITED LINES

Figure: Screenshot of a tab delimited .txt file and "\r\n" terminated line seen using notepad++

2. How to load a **NULL** value from the TEXT (tab delimited) file into the database via Load Infile command?

### Answer:

You would have to place **\N** in the TEXT file instead of leaving it blank. MySQL load infile command would treat **\N** as **null** when loading data into the database table.

3. While loading data using “LOAD DATA INFILE...” I get the error “The MYSQL server is running with the –secure-file-priv option so it cannot execute this statement”

**Answer:**

There are 2 alternatives to load the data if you get this error

- i. Use LOAD DATA **LOCAL INFILE** instead of **LOAD DATA INFILE**.
- ii. Place all data files in the folder path stated in the line of **mysql.ini** as shown below and then run the LOAD DATA command  
`secure_file_priv="c:/wamp64/tmp"`

Note: Running “show variables” command in SQL Workbench also shows you the value of the variable

4. Can I just ask the system to input certain lines?

**Answer:**

Yes. You can specify the starting line in LOAD DATA INFILE statement.

For example, the script

```
LOAD DATA INFILE 'C:/<folder>/GXTY/Data/order.txt' INTO TABLE order_t FIELDS  
TERMINATED BY '\t' LINES TERMINATED BY '\r\n' IGNORE 4 LINES;
```

This inserts the data from line 5 until the last line into the order table. This might be helpful if you want to append the data when there is an error during previous data loading and you have few lines of data already loaded.

5. Why do I find a special notation at the end of the data after loading the data into tables (like the one after ‘afae’)?

id	name
17	afae

**Answer:**

This is because on Windows typically two characters (carriage return/line feed) are used as a line terminator. So when you have not used the command phrase **lines terminated by '\r\n'** when loading data, this special notation gets inserted into the table.

Recommended script:

```
LOAD DATA INFILE 'C:/<folder>/GXTY/Data/order.txt' INTO TABLE order_t FIELDS  
TERMINATED BY '\t' LINES TERMINATED BY '\r\n';
```

6. *Why do I have quotes around columns of text data type in the data file?*

**Answer:**

We have given options to load data into the tables from .txt file. You could have also loaded data from .csv file (assuming .csv file contains data and the columns are delimited by ',' instead of tab). When creating the data file using Excel, Excel includes double quotes in the file around the contents of columns that has comma in them. Look at the data file using any Text Editor like Notepad to remove the quotes.

Or you can use the option **OPTIONALLY ENCLOSED BY ''** in the LOAD DATA statement as in

```
LOAD DATA INFILE 'C:/<folder>/GXTY/Data/order.csv' INTO TABLE order_t
FIELDS TERMINATED BY '\t' OPTIONALLY ENCLOSED BY '' LINES TERMINATED BY
'\r\n' IGNORE 1 LINES
```

7. *How to store a **BOOLEAN** (true/false) into a column?*

**Answer:**

MySQL creates the datatype **TINYINT(1)** for the column which is stated to be of **boolean** type. Note that a value of 0 is considered false; and non-zero values are considered true.

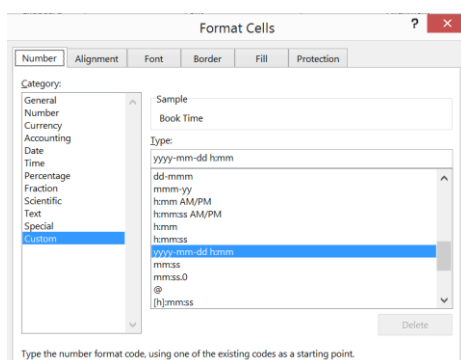
If you have a column which requires true/false type of values stored in it, you can choose to use CHAR(1) with 'T'/'F' inserted or to use TINYINT(1) with value 0 for false or 1 for true inserted.

## 8. Why is data for date columns entered as 0000-00-00?

### Answer:

For Load Data Infile to correctly input date columns, you have to ensure that the date in the given .txt file is in the format year-month-date either yyyy-mm-dd or yy-mm-dd because MySQL only recognizes this format as date and validates them before inserting.

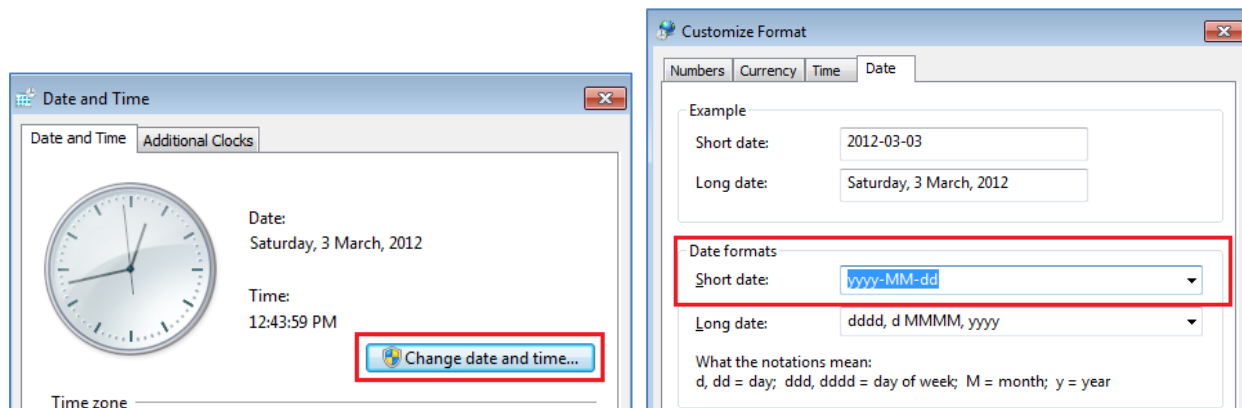
If you are given an excel file containing date columns, ensure they are in the format year-month-date. You could customize the column containing date by selecting the column, right click → Format Cells to change the format to the required format.



If you are unable to change the format of the “date” column, change the way Windows display date, datetime first before formatting, you can do the following:

On Windows, right click on the clock from the taskbar, select ‘Adjust Date/Time’ and click on ‘Change date and time’. In the new window that appears click on the link ‘Change Calendar Settings’ to change the date and time format. You can use the drop-down list in ‘Short date’ to select from formats available or can enter the needed format directly.

You can then use the column containing date in the excel file to follow the format.



The above steps are to ensure that when date is entered in Excel, it is in the required format.

Alternately, you can use MySQL function and indicate the format that the date type value is in the .txt file. The MYSQL function converts the date value in the proper format before inserting the data. Assume you are inserting data into Order\_t table that contains columns order\_id, and order\_date, and the .txt file contains date in the format dd/mm/yyyy, use

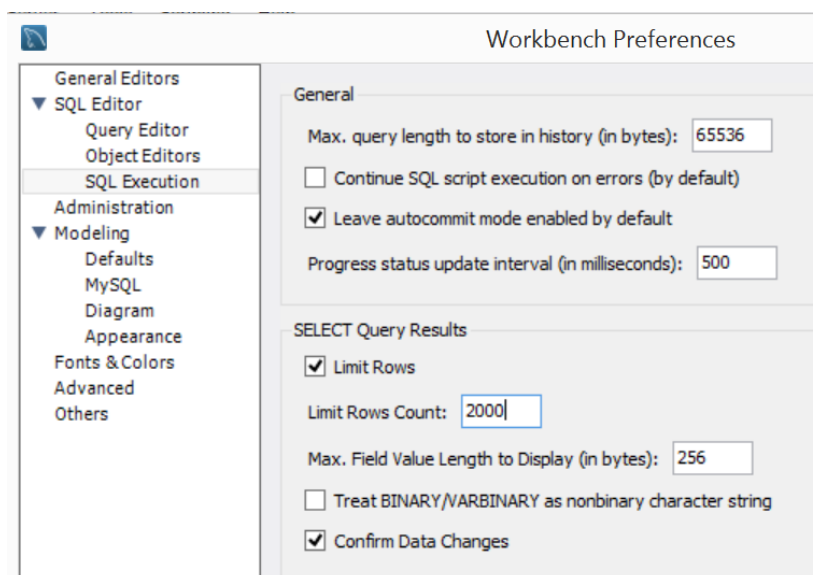
```
LOAD DATA INFILE 'C:/<folder>/GXTY/Data/order.txt' INTO TABLE order_t FIELDS
TERMINATED BY '\t' LINES TERMINATED BY '\r\n' ignore 1 lines (order_id, @order_date)
SET order_date = STR_TO_DATE(@order_date,'%d/%m/%Y');
```

(The order\_date after SET keyword is the name of the column in the table order\_t)

### 9. Why am I not able to view all records in the table?

#### Answer:

MySQL Workbench by default imposes the limit to display 1000 records. In order to remove the restriction, choose *Edit -> Preferences* from the menu. Select option SQL Execution (or SQL queries) on the left panel of the window Workbench Preferences. In the text box “*Limit Rows Count*” increase the limit to a suitable number or uncheck the box “*Limit Rows*”



~~ THE END ~~