

# Modelling with R and MySQL

## - Manual -

November 13, 2014

### Contents

<b>1</b>	<b>Installation of Software</b>	<b>2</b>
1.1	Database . . . . .	2
1.1.1	Linux: Apache-Server, MySQL-Datenbank and phpMyAdmin . . . . .	2
1.1.2	Windows: Xampp . . . . .	3
1.1.3	Mac: MAMP . . . . .	4
1.2	Import data into the database . . . . .	4
1.2.1	Import with phpMyAdmin . . . . .	4
1.2.2	Import via terminal . . . . .	5
1.3	R . . . . .	5
1.3.1	Linux . . . . .	5
1.3.2	Windows . . . . .	6
1.3.3	Mac . . . . .	6
1.4	RStudio . . . . .	6
1.5	RMySQL . . . . .	6
1.5.1	Linux . . . . .	6
1.5.2	Windows: install RMySQL with Rtools . . . . .	7
1.5.3	Mac . . . . .	7
<b>2</b>	<b>Server</b>	<b>8</b>
<b>3</b>	<b>Help Sources and Literature</b>	<b>8</b>
3.1	Database . . . . .	8
3.2	R . . . . .	8
3.2.1	Books . . . . .	8
3.2.2	Videos . . . . .	8
3.2.3	Forum . . . . .	9
3.2.4	Command Overviews and Fast help . . . . .	9

# 1 Installation of Software

## 1.1 Database

### 1.1.1 Linux: Apache-Server, MySQL-Datenbank and phpMyAdmin

<https://help.ubuntu.com/community/ApacheMySQLPHP>

On this page, different alternatives are described. The simplest way is to first install the package lamp apache (server), then mysql (database) and afterwards phpMyAdmin, which is a graphical user interface. In most cases, a small change has to be made in one script, so that phpMyAdmin can communicate with the apache server, which is called apache2. The detailed explanation of all steps follows:

Write down in the terminal:

```
> sudo apt-get install tasksel
> sudo apt-get update
> sudo tasksel install lamp-server
```

Password for the MySQL administrative:

```
> root
user:
> user123
```

Try the following line in your browser (for example Firefox): <http://localhost/>

If the following message appears:

“It works! This is the default web page for this server. The web server software is running but no content has been added, yet.”

you can install phpMyAdmin. It is important to install the MySQL-Server first before you can install phpMyAdmin afterwards.

```
> sudo apt-get install phpmyadmin
```

- During installation you have to choose the server. Choose **apache2**.
- Question “configure database for phpmyadmin with dbconfig-common?” Click **yes**.
- “Please provide the password for the administrative account with which this package should create its MySQL Database and user:”  
> user123
- “Password for the databases administrative user:”  
> user123

- “MySQL” application password for phpmyadmin”  
`> user123`
- if you have tried accidentally to install phpMyAdmin first, before you installed php-MyAdmin, you have to remove phpMyAdmin by the following command:  
`> apt-get autoremove phpmyadmin`  
 and then start again
- If phpMyAdmin seems to be installed test it if the following line in the browser works:  
`http://localhost/phpmyadmin`
- Yes? Congratulations!
- No? Then write in a terminal:  
`> sudo gedit /etc/apache2/apache2.conf`
- An editor opens. You have to paste the following line at the end of the page , save the document and then close it:  
`Include /etc/phpmyadmin/apache.conf`
- Then you have to restart the server by writing into the terminal:  
`sudo /etc/init.d/apache2 restart`
- `http://localhost/phpmyadmin` in your browser should open your database

With the username `root` and the password `user123`, you are now able to enter you local database

### 1.1.2 Windows: Xampp

Download xampp from <http://www.apachefriends.org/en/xampp-windows.html>

You can choose if you want Apache and MySQL to be started automatically when Windows is started. Then you have to set the checkmarks when the question is raised if they should be installed as a service. Otherwise you have to start Apache in MySQL manually when required. All other default settings are okay.

You can start the services manually with the Xampp Control Panel. This is stored in `C:\xampp` and opens by clicking on `xampp-control.exe`. Then you have to click the start button of Apache and MySQL. After usage, you have to click on the Stop buttons and close the Xampp Control Panel by exit.

You can administer the database with phpMyAdmin. To open phpMyAdmin, you have to call the following line in your browser:

`http://localhost/phpmyadmin`

If the message “Server not found“ appears, you have to click in the menu **File > Work offline**. If a red error message appears, quite often it helps to empty the cache. For that click, in the menu **Tools > Clear recent history**. Afterwards you can then see the databases. To be able to run the same code pieces, we change the user rights.

For that, on the main page <http://localhost/phpmyadmin/> go to the tab Privileges (Rechte).

Change the rights of the user **root** for the host **localhost** with the icon on the rights. Activate password in the third box named **Change password** and insert **user123** .

If problems with PhpMyAdmin appear, you have to change the configuration: Open the file **config.inc.php**, which is located in the folder **C:\Xampp\PhpMyadmin** with a texteditor (If you click on pagebreak in format in the texteditor, it looks better). Below **Authentication type** and **info** the following should be written:

```
\\$cfg['Servers'][$i]['auth_type']      = 'cookie';
\\$cfg['Servers'][$i]['user']          = 'root';
\\$cfg['Servers'][$i]['password']      = '';
\\$cfg['Servers'][$i]['AllowNoPassword'] = true;
```

Then, for this application, cookies have to be accepted from localhost. The password will be asked when opening the database.

### 1.1.3 Mac: MAMP

For the installation under Mac OS, just download MAMP from <http://www.mamp.info> and follow the instructions. If you are asked to create a user choose **root** with the password **user123**.

To start the Apache and MySQL Server just start MAMP under **Applications**. A new window opens and the server should start by default (= green light). With **Open start page** you can open the phpMyAdmin-Interface in your browser.

## 1.2 Import data into the database

There are different ways of getting data into the database. Most of them can be utilized by terminal or by the graphical user interface phpMyAdmin. phpMyAdmin has a maximum size limit for files by default which could be changed. So if you want to import large data sets it is better to use the terminal. In the following, the import of the file formats csv and sql is explained.

### 1.2.1 Import with phpMyAdmin

If the new data should go into a new database, the first thing you have to do is to create a database. This can be done on the main page of phpMyAdmin.

**csv** csv file consist of data in a table. The table structure should already be prepared in the database. If the csv file has column headings, the format / table structure can be

automatically detected, but if you want to determine the format of the columns you should make a table structure before importing the data. Then, when you are in the right database and the right table, click tab import. Choose csv as file format and browse the csv-file you want to import. The database shows a mistake if the data is not consistent with the table structure. This method is rather suitable for testing, change little datasets or if you do not have a choice because the data is just available in this format.

**sql** From a table or database export of an existing database, you get sql files (“sql dump”). They are written in the database language SQL. They contain information about the tables and the table structure additionally to the data. Because of that, you do not have to determine the table structure in advance. You can import more tables and even more databases with one sql file. In advance you have to choose - or if it doesn’t exist yet - create - the right name of the database. If you just import tables and there is no database name stored in the sql-file you can choose the database you want. Choose the tab **import**.

### 1.2.2 Import via terminal

Set the directory to the folder where the file you want to import is located. Then import the file by giving the username, password, database and file name. Here the example folder is called `Documents/renpass/database_files`, the example database is called `pathways` and the file is called `solar_120331.sql`. You can get out of the folder by just typing `cd` and then enter.

For **Linux / Mac**:

```
> cd Documents/renpass/database_files
> mysql -uroot -puser123 pathways < solar_130331.sql;
```

For **Windows**:

```
> cd C:\Documents\renpass\database_files
> C:\xampp\xampp\mysql\bin\mysql -uroot -puser123 pathways < solar_130331.sql;
```

## 1.3 R

### 1.3.1 Linux

```
> sudo apt-get update
> sudo apt-get install r-base
> sudo apt-get install r-base-dev
```

=> now it is possible to use the command `install.packages()` in R.

### 1.3.2 Windows

Download and install the latest R Version from

<http://cran.r-project.org/bin/windows/base/>.

The standard settings are okay. Both versions, 32-bit and 64-bit will be installed.

### 1.3.3 Mac

Download and install the latest R Version from

<http://cran.r-project.org/bin/macosx/>

## 1.4 RStudio

Download the appropriate version for your system from

<http://www.rstudio.com/ide/download/desktop> and install it by a doubleclick.

RStudio is a very handy user interface for R. Editor and R console, plots, and help are in one overview. In the upper left window is the editor, where you can work on code files and save them. Different colours help to identify commands and structure in the code. If you start a new script you can choose it to be handled like an R script by choosing R script in the right corner at the bottom of the editor area. Then you can directly see the R syntax highlighting. In the upper right corner, you have three possibilities to send code to R:

- **Run** hands over the marked area or the line in which the cursor is to R. This can be one command or the whole document or anything in between. You can see it appear in the R area below the editor area.
- The icon next to it can be used to run the same selection again. Any changes to the selection are considered in this re-run.
- **Source** commits the whole file to R. In this case, the code is not copied into the R window.

Windows:

You have to check if the correct R Version is connected to RStudio. In **Tools > Options > R general** the version can be changed. This should be 32-bit version of your R installation.

## 1.5 RMySQL

RMySQL is a R package and enables the direct database connection.

### 1.5.1 Linux

Terminal:

```
> sudo apt-get install r-cran-rmysql
```

For testing if this was successful, you can open R in the Terminal by writing

```
> R
```

in the Terminal and then type

```
> require(RMySQL)
```

### 1.5.2 Windows: install RMySQL with Rtools

- Download Rtools from <http://cran.r-project.org/bin/windows/Rtools/>. When asked for a path during installation, select the 32bit bin folder of your R installation. This will look similar to this: `C:/Programme/R/R-VERSIONNUMBER/bin/i386/`.
- When asked if environment variables or system path should be added, click yes! Otherwise this has to be done manually afterwards.
- Open a text editor as administrator (right-click on **Start** > **Programme** > **Zubehör** > **Editor** > **Als Administrator ausführen**).  
Type  
`MYSQL_HOME=C:/xampp/mysql`  
The path must be exactly the path where the mysql folder is on your computer. Make sure to use slash (/) and **no** back slash (\).  
Save the file to `C:/Program Files/R/R-VERSIONNUMBER/etc/i386` as `Renviron.site`. `Renviron` is the name of the file and `.site` is the file extension. Be careful that the file is not automatically saved as a `.txt` file.
- If it does not exist, create a folder named `opt` in `C:/xampp/mysql/lib/` and copy all the loose files from the `lib` folder to `opt`.
- Copy the file `libmysql.dll` from the `opt` folder and paste it to  
`C:/Programme/R/R-VERSIONNUMBER/bin/i386/`  
Also paste it to  
`C:/xampp/mysql/bin/`
- Type `install.packages('RMySQL',type='source')` in your R console. Allow the system to create a library.

Warnings after installation can usually be ignored but there should be no error. When the installation was successful, the RMySQL package can be loaded by typing

```
> require(RMySQL)
```

to the R console.

Help installing RMySQL in Windows could also be found at

<http://www.ahschulz.de/2013/07/23/installing-rmysql-under-windows/>

### 1.5.3 Mac

Type in your R console

```
> install.packages('RMySQL',type = 'source')
```

Allow the system to create a library.

## 2 Server

We use a server for the course to have a platform for communication and exchange of data. You can find the main page for the course at

`https://renpass.eu/moodle`

Moodle is an environment for e-learning where you can find all important dates, additional information, voluntary quizzes, a forum and more. For the first login use your first name and the password `user123`. You should change your password after the first login under `Settings > My profile settings > Change password`. Your profile can be edited under `Settings > My profile settings > Edit profile`. There you can as well choose your preferred language (English or German).

## 3 Help Sources and Literature

### 3.1 Database

`http://dev.mysql.com/doc/`  
`http://www.1keydata.com/de/sql/sql-select.php`  
`http://www.w3schools.com/sql/default.asp`

### 3.2 R

#### 3.2.1 Books

Books for getting into R by reading or to find answers to special problems by searching keywords in them.

- Uwe Ligges. *Programmieren mit R*. Springer, 2 edition, 2007
- Winston Chang. *R Graphics Cookbook - Practical Recipes for Visualizing Data*. O'Reilly Media, 1 edition, December 2012. URL `http://www.cookbook-r.com/`
- Michael J. Crawley. *The R Book*. John Wiley and Sons, Ltd, 2007

#### 3.2.2 Videos

An R course that is split into several video lectures is available on `http://www.youtube.com/user/rdpeng` > `Playlists > Computing for Data Analysis`.



### 3.2.3 Forum

- The Cran R Page: <http://cran.r-project.org/>
- An R Wiki: <http://rwiki.sciviews.org/doku.php>
- A useful forum where a lot of R problems are discussed: <http://stackoverflow.com/tags/R>

### 3.2.4 Command Overviews and Fast help

- Four pages overview of common functions, very helpful:  
Tom Short. *R Reference Card*. EPRI PEAC, tshort@epri-peac.com, 11 2004. URL <http://cran.r-project.org/doc/contrib/Short-refcard.pdf>
- If you want to have a function explained, for example the function `mean()` just type `?mean` and in the lower right corner of RStudio an explanation and examples as well as similar functions appear.

The R community is so huge that almost every problem that can occur during coding or using code from R has happened to somebody else already. So don't hesitate to just search in the big internet for your error message, your warning, your function, or your problem. You will find more information if you search in English. It is often useful to search for "R" + "statistical" + "your search words" to make sure you get the right R in your results.