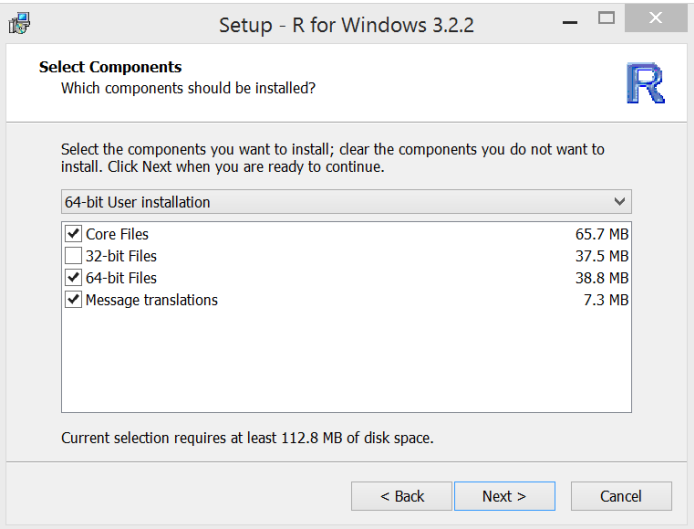
SMOKE-MOVES Input QA Tool Manual

This manual provides step-by-step instructions for running the SMOKE-MOVES Input QA Tool. The tool will produce and store the full set of QA reports on the user's local machine for any annual version of the SMOKE-MOVES input databases.

Download R

- https://www.r-project.org/
- Go to the R programming website and select **download R**
- Choose a CRAN Mirror close to your location to download the program files from.
- Choose the correct version for your operating system.
- Select **install R for the first time**
- Select **Download R.3.2.2 for Windows/Linux/OS** (versions will change as R is updated)



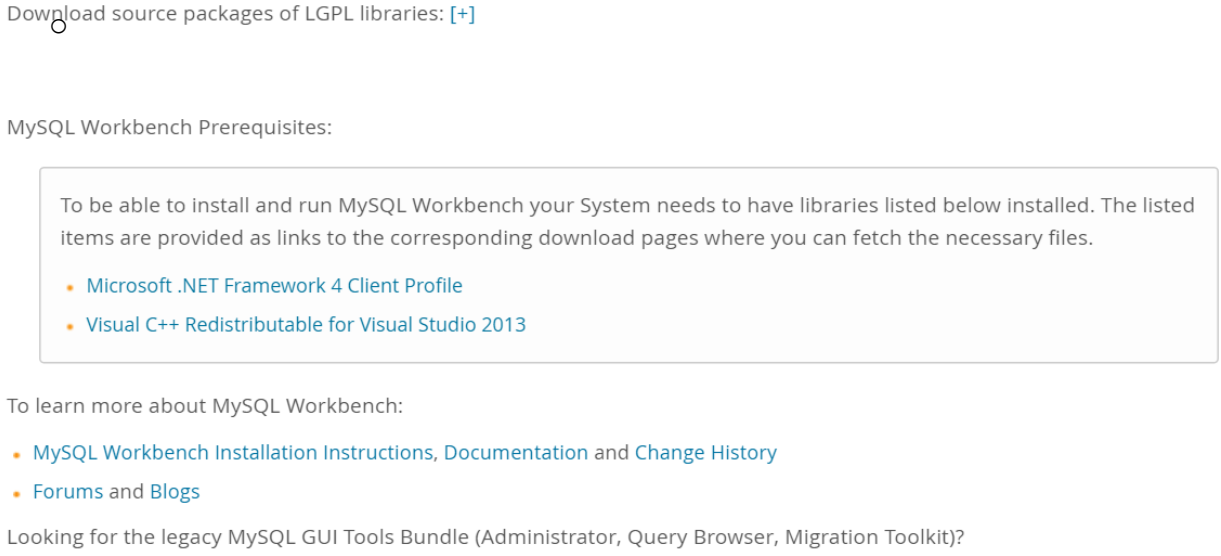
- Choose the correct components for your operating system

Download R user interface – Rstudio

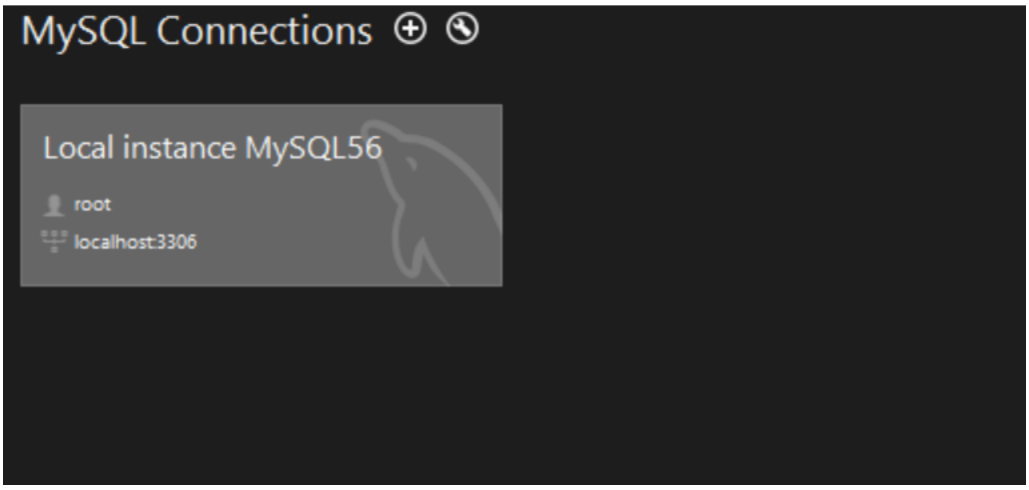
- https://www.rstudio.com/products/rstudio/download/
- Download the correct version for your operating system

Querying MySQL Data

- MySQL databases need to be compiled in Rstudio.

- Install RMySQL server: http://dev.mysql.com/tech-resources/articles/mysql-installer-for-windows.html
    - Remember password specified during installation
- Find MySQL Workbench download and install:
    - Once you reach the installation page, pay attention to any prerequisite programs listed above the download (as seen below)
    - 

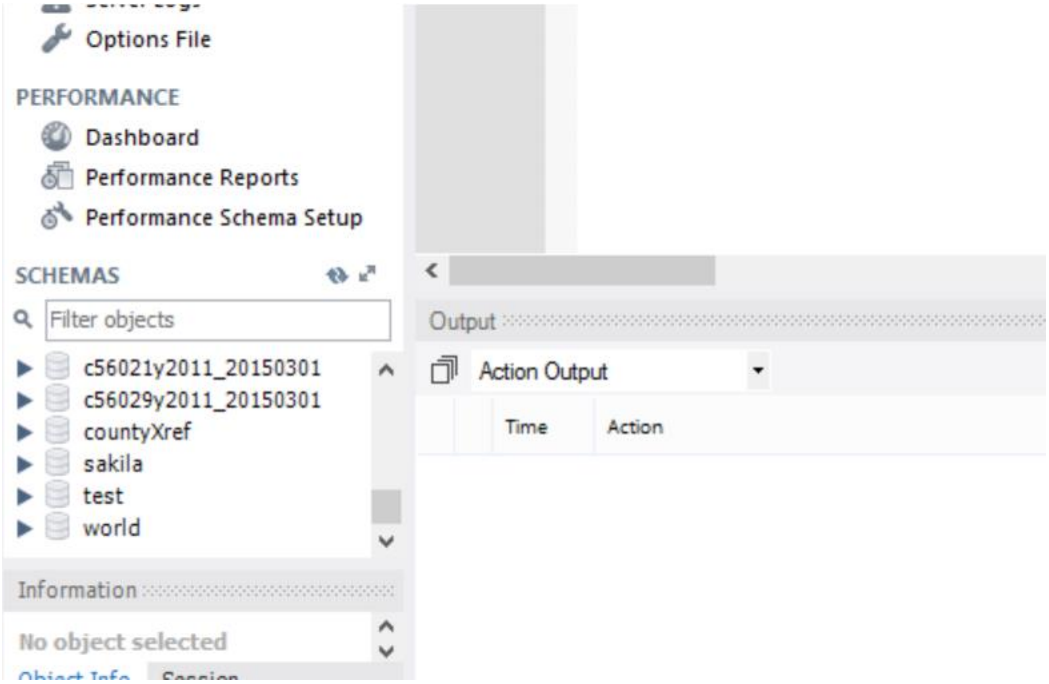
- Configure MySQL workbench with system
    - Open workbench. The interface seen below will show up. Click grey box, login. (see picture below)



- Login will be username and password set upon installation of the MySQL server. (not workbench) "root" is the default username unless otherwise specified.
- Place .myi,.myd,.frm files (downloaded MOVES files) in the directory for MySQL workbench. Individual counties must be represented as folders (should already be in this format) with .myi,.myd,.frm files inside. The folders must be placed in the default

MySQL server directory: “C:\ProgramData\MySQL\MySQL Server 5.6\data” (may vary by machine)

- Note that C:\ProgramData is a hidden directory that must be searched in the Documents window to be found. It will not show up in the C: directory. The MySQL server directory may vary by computer, but the county folders must be saved in the data folder of MySQL Server.



- If you experience difficulty finding the directory, MySQL should have a few pre-saved files in the server directory, labeled as schemas in the workbench. These files (and their directory) can be located with a documents search for “sakila”, test”, etc. The file names above “sakila” in the screenshot are the files for counties 56021, and 56029, which should show up in the workbench schemas window after you’ve placed them in the server directory.

Problem MySQL Data

Some of the datatypes don’t run properly or are empty. These dataframes are not used for any of the graphics.

No data in the table

- Auditlog
- Driveschedulesecondlink
- Fuelsupply
- Importstartsopmodedistibution
- Link

CountyID won't attach to table

- Converttempmessages
- Road type

CountyID won't attach to table AND no data

- Linksourcetypehour
- Onroadretrofit
- Opmodedistribution
- Startshourfraction
- Startsmonthadjust
- Startssourcetypefraction
- Zonemonthhour

Data is returned in a unusual format

- fuelformulation

.CSV files to Rstudio

Comma separated files can be converted to .Rdata files in Rstudio by using the read.csv command with the filepath.

X<-read.csv(file=(filepath with \\ instead of \) , header=(TRUE/FALSE depending on column headers) , sep="(, .or . or / etc.)")

The commands for loading all the necessary .csv files are inside of the fileload masterscript. However, the user must place the .csv files in the correct directory for the script to find. In the middle of the directory masterscript command, the user will be prompted to place all of the SMOKE, countylist, and countyxref files in the **smoke** directory. This directory will be specified in the script.

R Masterscripts

Loading R packages: load.library.R masterscript

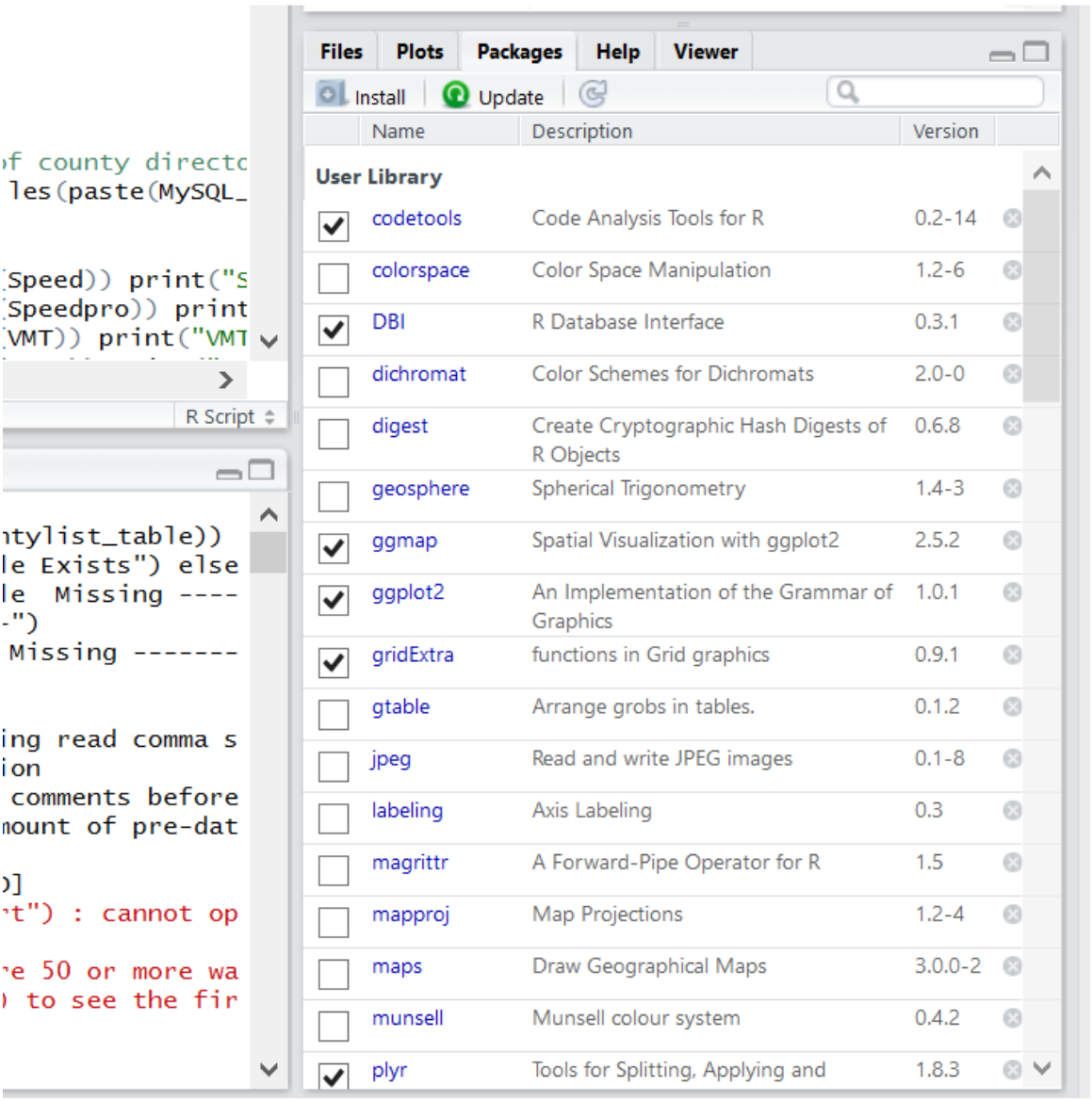
A series of packages are needed to run the script. In order to avoid issues with the script becoming outdated due to the evolution of package versions, the following version specific packages must be installed:

Package	Version
codetools	0.2-14
DBI	0.3.1
ggmap	2.5.2
ggplot2	1.0.1
gridExtra	0.9.1
plyr	1.8.3
RMySQL	0.10.7

The load.library package masterscript should install the correct packages on the user's machine. However, Rstudio has issues reading package installation script and the user may need to manually install the packages.

Run the load.library.R script.

The correct version of the packages should show up under the **User Library** in the packages tab of the lower right window in Rstudio. The box next to the package should be checked, signifying its activation.



If the correct packages aren't installed, the user needs to install the packages manually. If the incorrect version is installed, the user must uninstall the packages with the command remove.packages(). The package name without quotation marks must be typed in the parenthesis. The user must then install the packages by using the command install.packages() with the package URL saved as a variable in the same format as seen in load.library.R. The URL and sample command for DBI package can be seen below:

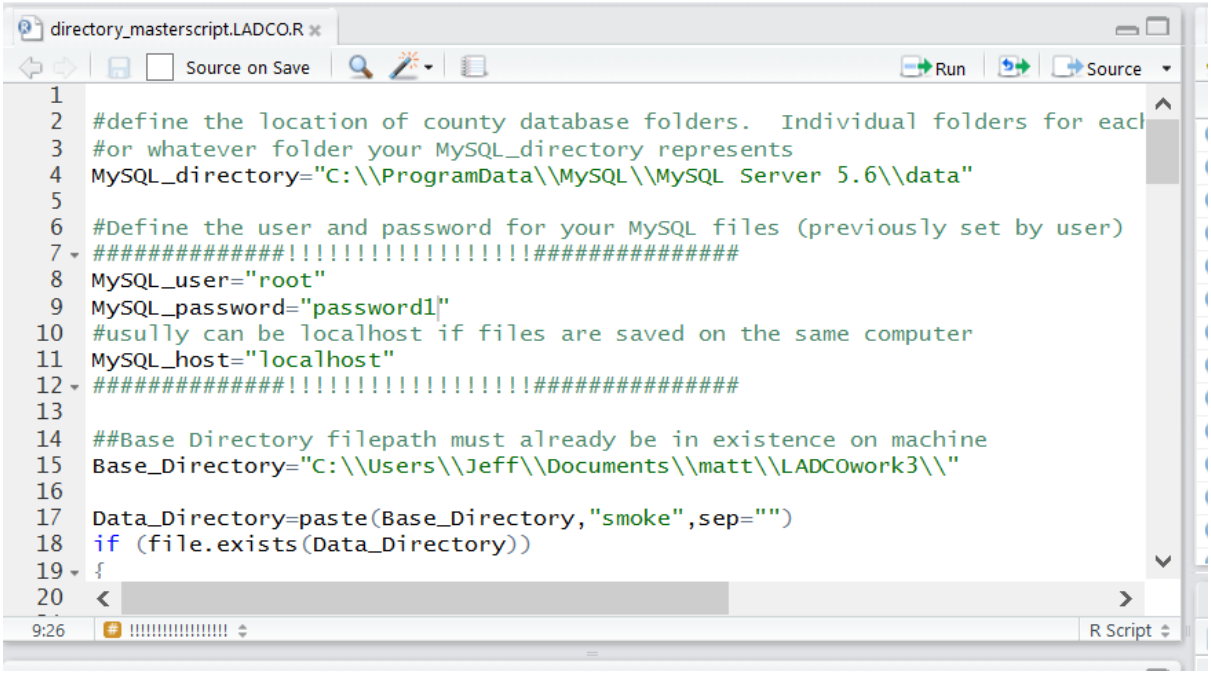
```
packageurl<-"https://cran.rstudio.com/bin/windows/contrib/3.2/DBI_0.3.1.zip"
install.packages(packageurl, repos=NULL, type="source")
```

The user must make sure the box is checked next to each package.

Directory Creation: directory_masterscript.LADCO.R

- The directory masterscript creates the directory for all of the outputs as well as function scripts.
- Before directory_masterscript is run the variables must be set inside the script:

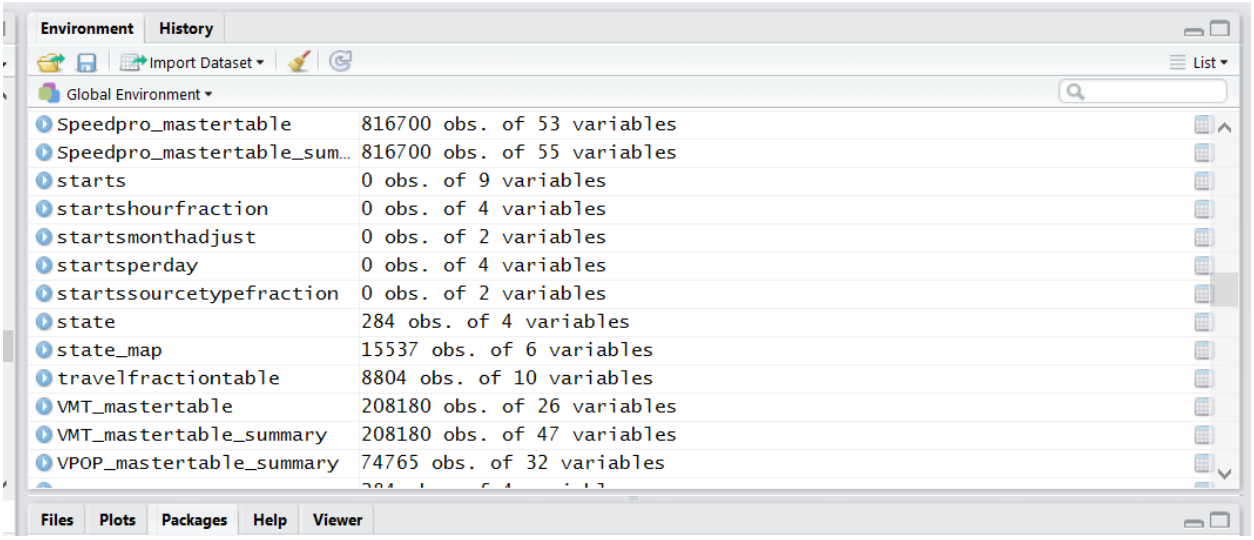
Variable	Location
MySQL_directory	Location of SQL files. Usually the "~ProgramData\\MySQL\\MySQL Server 5.6\\Data". The files in MySQL workbench can be searched as specified earlier.
MySQL_user	Default is "root" unless specified otherwise by user during installation of RMySQL Server
MySQL_password	Set by user during RMySQL server installation
MySQL_host	"localhost" if R and SQL files are on the same machine
Base_Directory	The base directory for all of the simulation outputs and files. This directory must be created manually by the user and be in existence before directory_masterscipt.LADCO.R is run. The script builds off of this base_directory by adding a folder heirarchy inside



- Directories must have double slashes ("\\") separating folders and be inside of quotation marks
- In the middle of the directory masterscript, the script will pause and prompt the user to place the SMOKE, countylist, and countyxref files in the **smoke** directory. The directory script will have created a **"smoke"** folder inside of the base directory. Once the files have been placed there the user will press enter.

Data Upload: Database_Load_Final.R

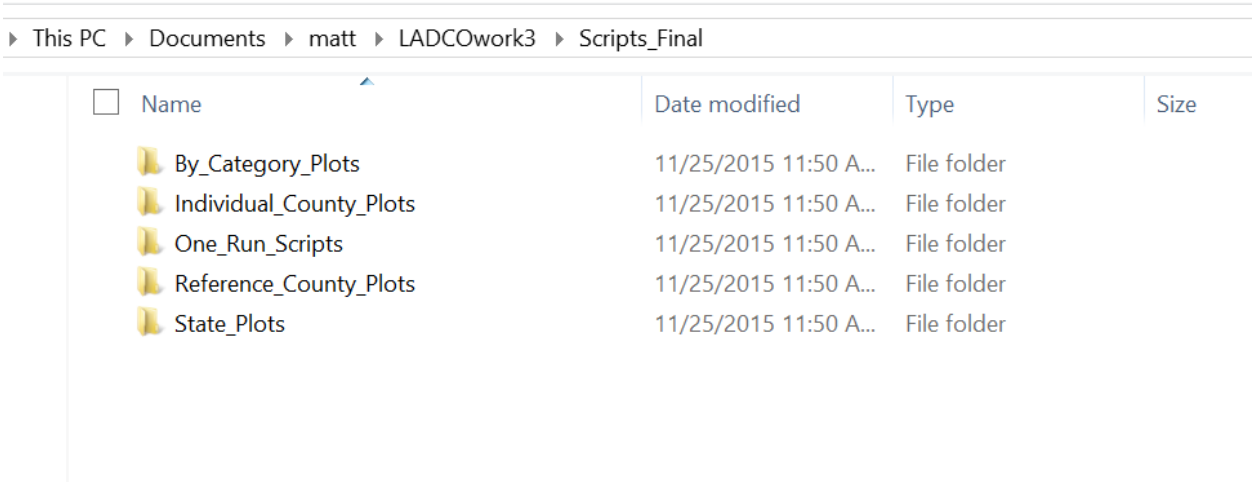
This script will convert all of the SQL and .csv files into .Rdata files and bring them into R. No user input is required. The script will take a while to run. Once it is finished, it is a good idea to save the environment created by the script. This way the user can reload the variable environment instead of rerunning the load data masterscript.



- The save icon in the upper left corner of the upper right screen of Rstudio will save the data environment.

Producing The Plots: Final_Plots_Masterscript.R

In order for the masterscript to find the plot functions in the final_plots_masterscript, the functions must be placed in the correct folders. The plot functions must be placed in the corresponding final_scripts folders created inside of your base_directory by directory_masterscript.R.



Some plots print an image to the lower right hand window in Rstudio and saves that .png file. In order to ensure that the graphics are of sufficient size, expand the lower right window to fill a large part of the screen.

Once the plot functions are placed in the folders Final_Plots_Masterscript.R can be run. It will take a long time and send the files to the “outputs” folder.

Some plots print an image to the lower right hand window in Rstudio and save that .png file. In order to ensure that the graphics are of sufficient size, expand the lower right window to fill a large part of the screen.