PRINCETON UNIVERSITY

Introduction to RStudio

(*v 1.3*)

Oscar Torres-Reyna

otorres@princeton.edu

August 2013

http://dss.princeton.edu/training/

Introduction

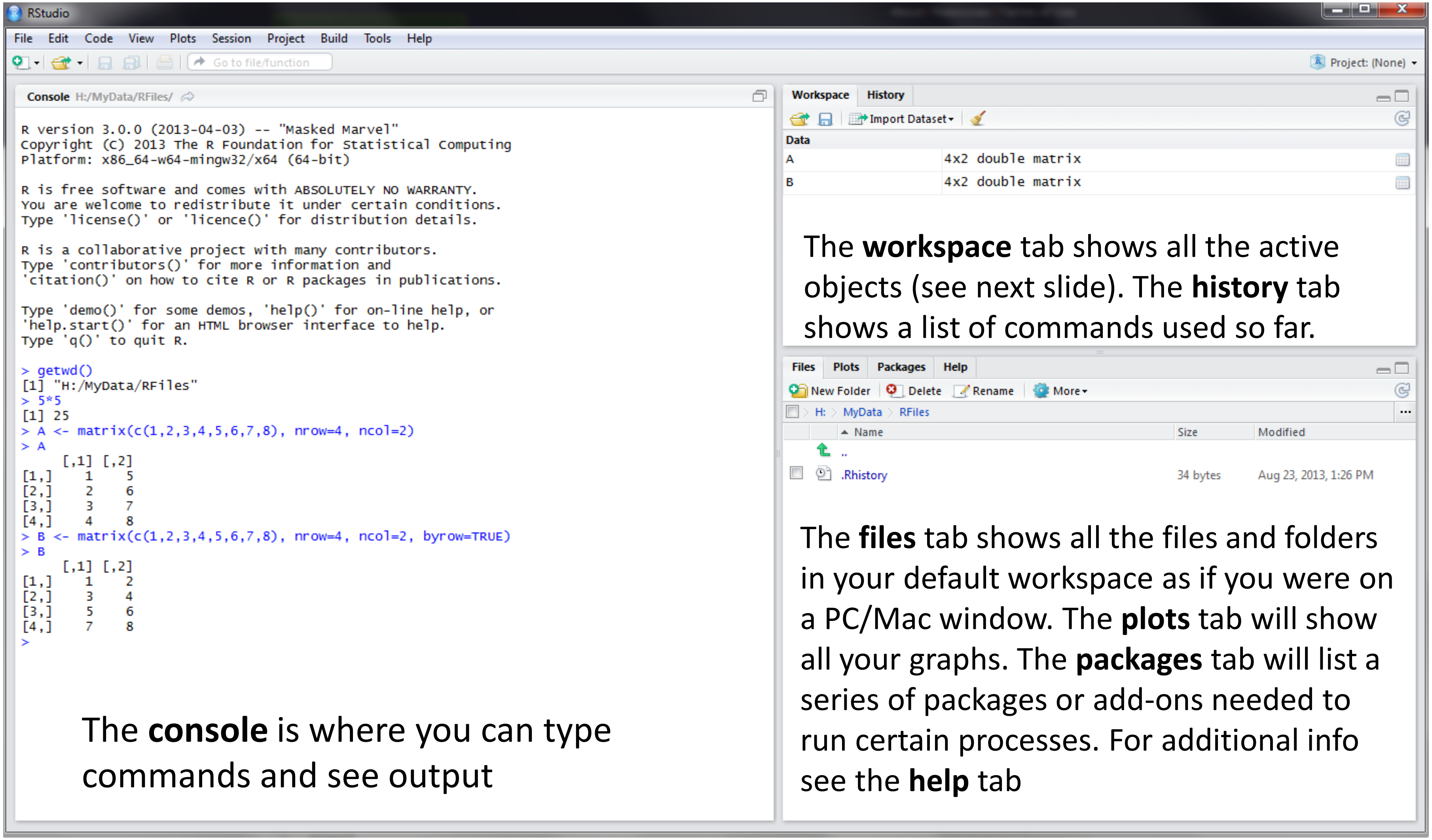
RStudio allows the user to run R in a more user-friendly environment. It is open-source (i.e. free) and available at http://www.rstudio.com/

For R related tutorials and/or resources see the following links:

http://dss.princeton.edu/training/

http://libguides.princeton.edu/dss

RStudio screen


RStudio
File Edit Code View Plots Session Project Build Tools Help
Go to file/function
Project: (None)
Console H:/MyData/RFiles/
Workspace History
Import Dataset
Files Plots Packages Help
New Folder Delete Rename More
H: MyData RFiles
Name Size Modified

```
R version 3.0.0 (2013-04-03) -- "Masked Marvel"
Copyright (C) 2013 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> getwd()
[1] "H:/MyData/RFiles"
> 5*5
[1] 25
> A <- matrix(c(1,2,3,4,5,6,7,8), nrow=4, ncol=2)
> A
     [,1] [,2]
[1,]    1    5
[2,]    2    6
[3,]    3    7
[4,]    4    8
> B <- matrix(c(1,2,3,4,5,6,7,8), nrow=4, ncol=2, byrow=TRUE)
> B
     [,1] [,2]
[1,]    1    2
[2,]    3    4
[3,]    5    6
[4,]    7    8
>
```

The **console** is where you can type commands and see output

Data

A	4x2 double matrix
B	4x2 double matrix

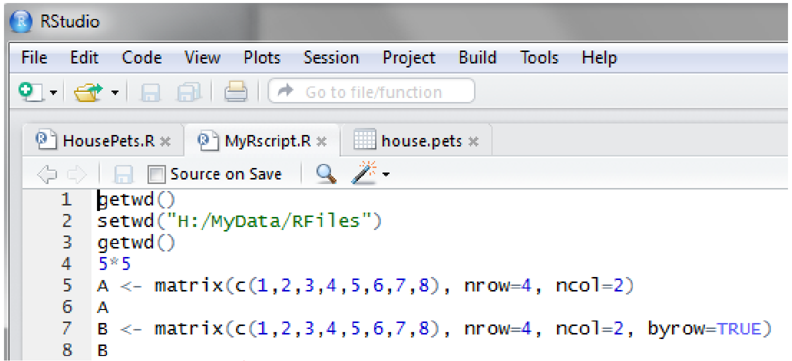
The **workspace** tab shows all the active objects (see next slide). The **history** tab shows a list of commands used so far.

Name	Size	Modified
..		
.Rhistory	34 bytes	Aug 23, 2013, 1:26 PM

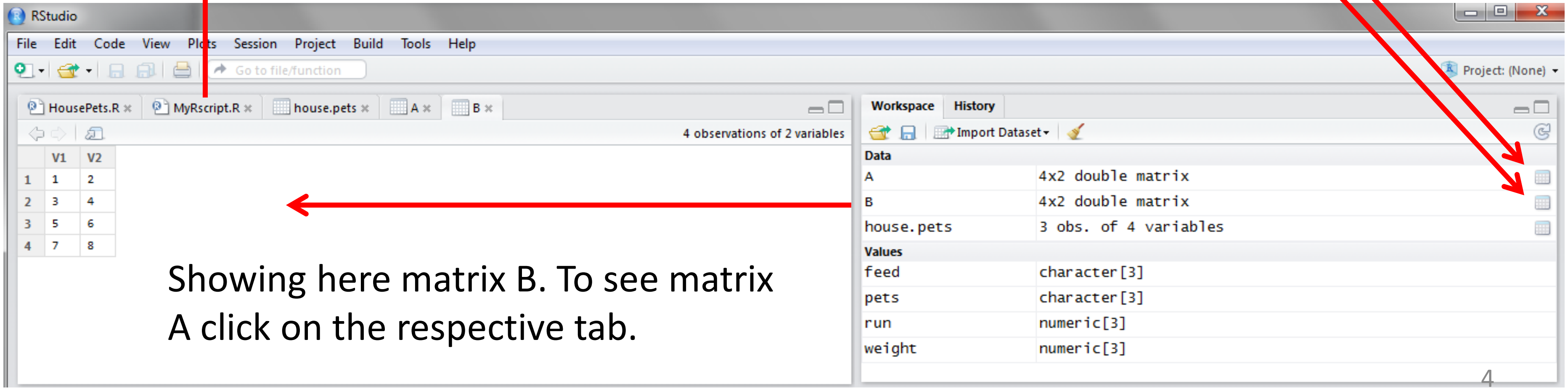
The **files** tab shows all the files and folders in your default workspace as if you were on a PC/Mac window. The **plots** tab will show all your graphs. The **packages** tab will list a series of packages or add-ons needed to run certain processes. For additional info see the **help** tab

Workspace tab (1)

The workspace tab stores any object, value, function or anything you create during your R session. In the example below, if you click on the dotted squares you can see the data on a screen to the left.



```
getwd()
setwd("H:/MyData/RFiles")
getwd()
5*5
A <- matrix(c(1,2,3,4,5,6,7,8), nrow=4, ncol=2)
A
B <- matrix(c(1,2,3,4,5,6,7,8), nrow=4, ncol=2, byrow=TRUE)
B
```



4 observations of 2 variables

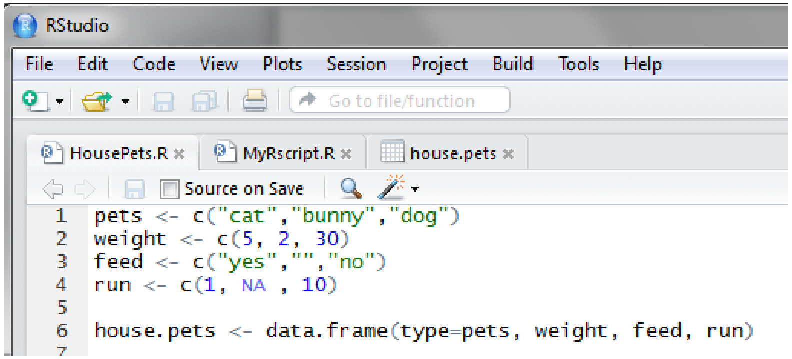
	V1	V2
1	1	2
2	3	4
3	5	6
4	7	8

Data	
A	4x2 double matrix
B	4x2 double matrix
house.pets	3 obs. of 4 variables
Values	
feed	character[3]
pets	character[3]
run	numeric[3]
weight	numeric[3]

Showing here matrix B. To see matrix A click on the respective tab.

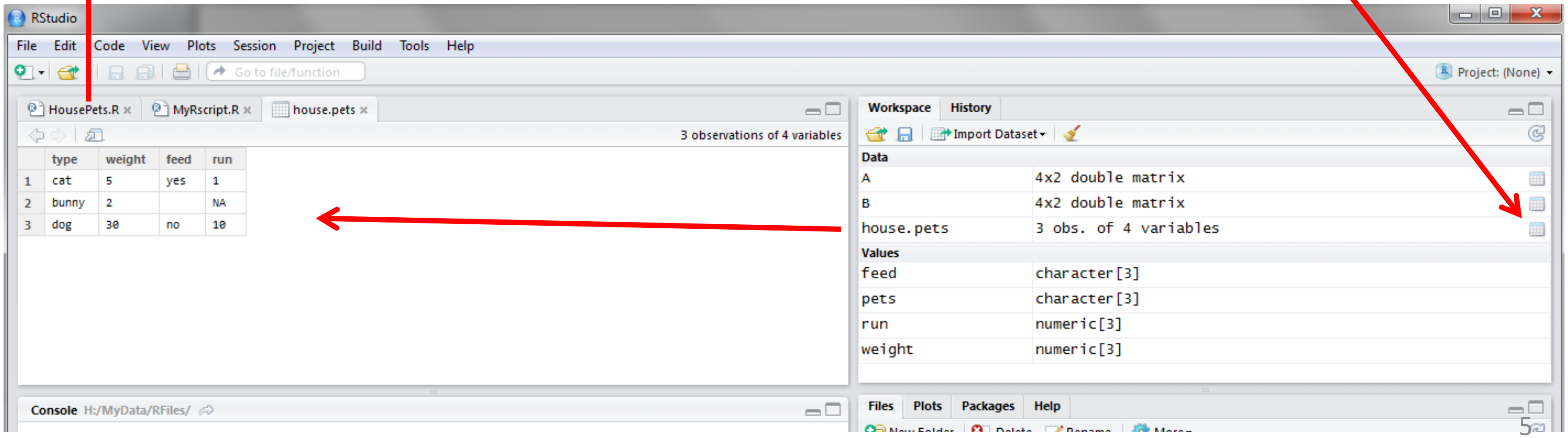
Workspace tab (2)

Here is another example on how the workspace looks like when more objects are added. Notice that the data frame `house.pets` is formed from different individual values or vectors.



```
pets <- c("cat","bunny","dog")
weight <- c(5, 2, 30)
feed <- c("yes","","no")
run <- c(1, NA , 10)

house.pets <- data.frame(type=pets, weight, feed, run)
```

Click on the dotted square to look at the dataset in a spreadsheet form.


RStudio
File Edit Code View Plots Session Project Build Tools Help
Go to file/function
Project: (None)
HousePets.R
MyRscript.R
house.pets
Workspace
History
Import Dataset
Console H:/MyData/RFiles/
Files Plots Packages Help
New Folder Delete Rename More

3 observations of 4 variables

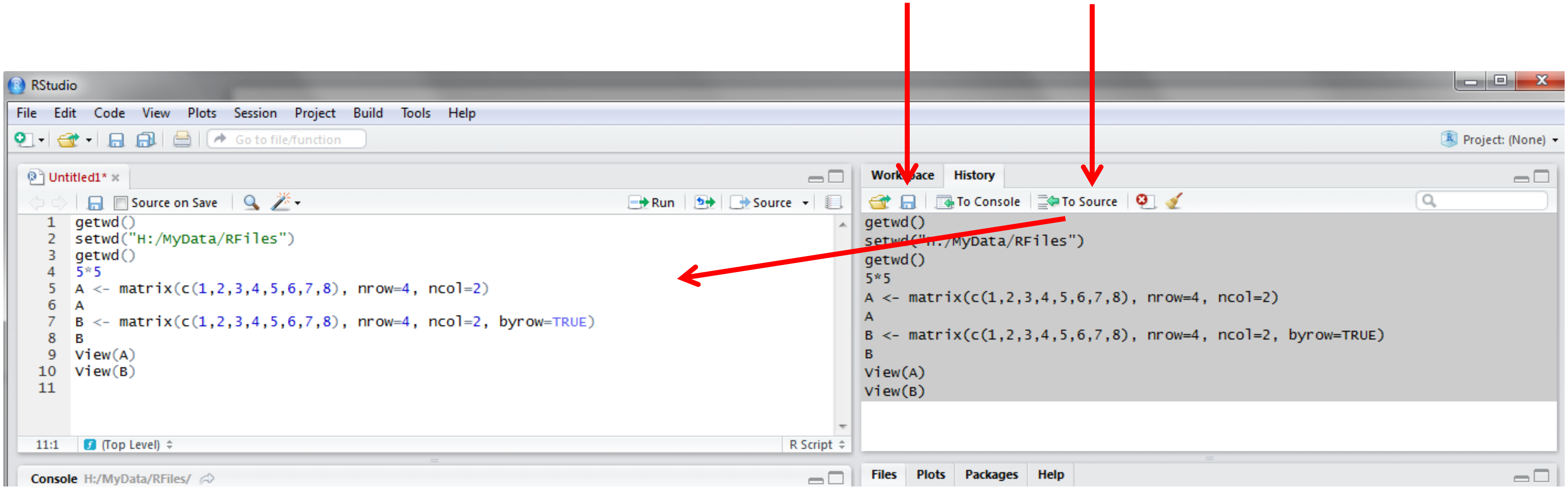
	type	weight	feed	run
1	cat	5	yes	1
2	bunny	2		NA
3	dog	30	no	10

```
Data
A            4x2 double matrix
B            4x2 double matrix
house.pets   3 obs. of 4 variables
Values
feed         character[3]
pets         character[3]
run          numeric[3]
weight       numeric[3]
```

History tab

The history tab keeps a record of all previous commands. It helps when testing and running processes. Here you can either **save** the whole list or you can **select** the commands you want and send them to an R script to keep track of your work.

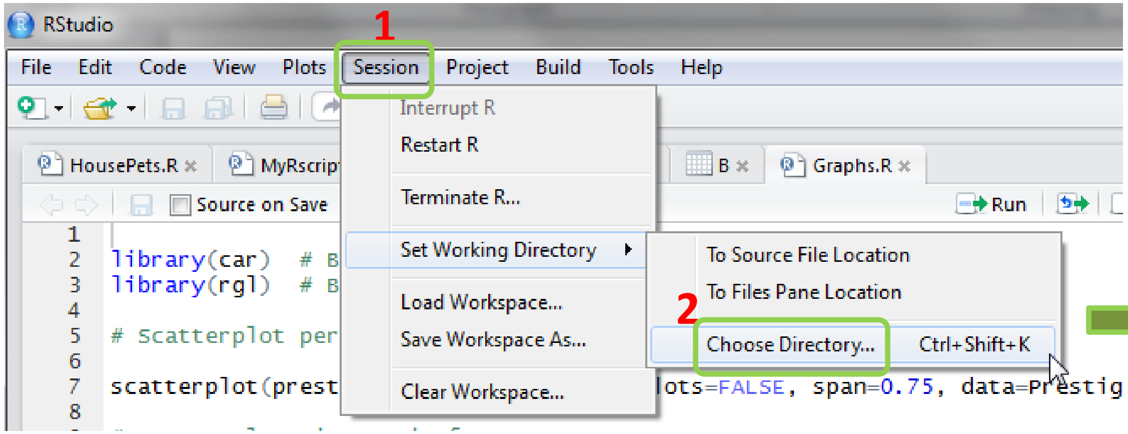
In this example, we select all and click on the “To Source” icon, a window on the left will open with the list of commands. Make sure to save the ‘untitled1’ file as an *.R script.



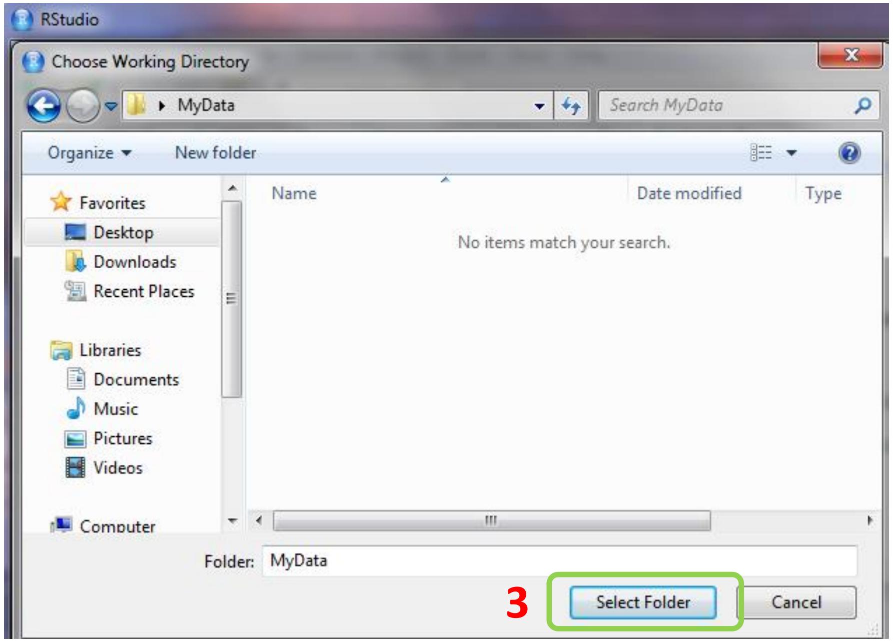
```
getwd()
setwd("H:/MyData/RFiles")
getwd()
5*5
A <- matrix(c(1,2,3,4,5,6,7,8), nrow=4, ncol=2)
A
B <- matrix(c(1,2,3,4,5,6,7,8), nrow=4, ncol=2, byrow=TRUE)
B
View(A)
View(B)
```

```
getwd()
setwd("H:/MyData/RFiles")
getwd()
5*5
A <- matrix(c(1,2,3,4,5,6,7,8), nrow=4, ncol=2)
A
B <- matrix(c(1,2,3,4,5,6,7,8), nrow=4, ncol=2, byrow=TRUE)
B
View(A)
View(B)
```

Changing the working directory



```
library(car)  # B
library(rgl)  # B

# Scatterplot per

scatterplot(prest                    lots=FALSE, span=0.75, data=Prestig
```



No items match your search.

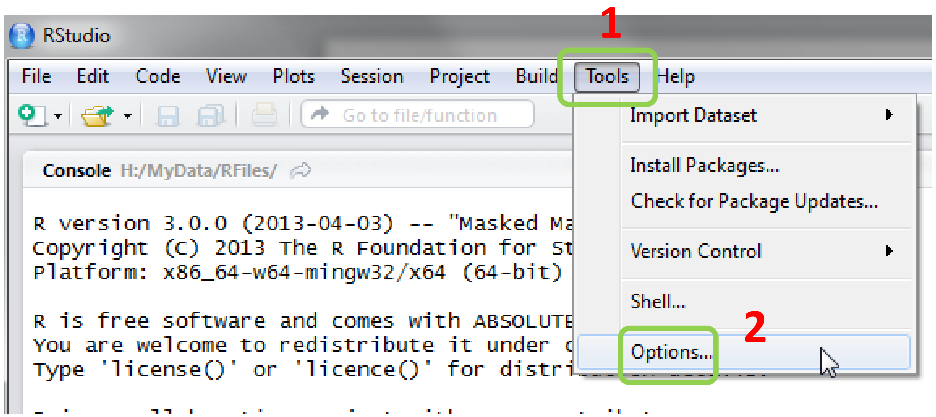
Folder: MyData

If you have different projects you can change the working directory for that session, see above. Or you can type:

```
# Shows the working directory (wd)

getwd()

# Changes the wd

setwd("C:/myfolder/data")
```

More info see the following document:
http://dss.princeton.edu/training/RStata.pdf

Setting a default working directory

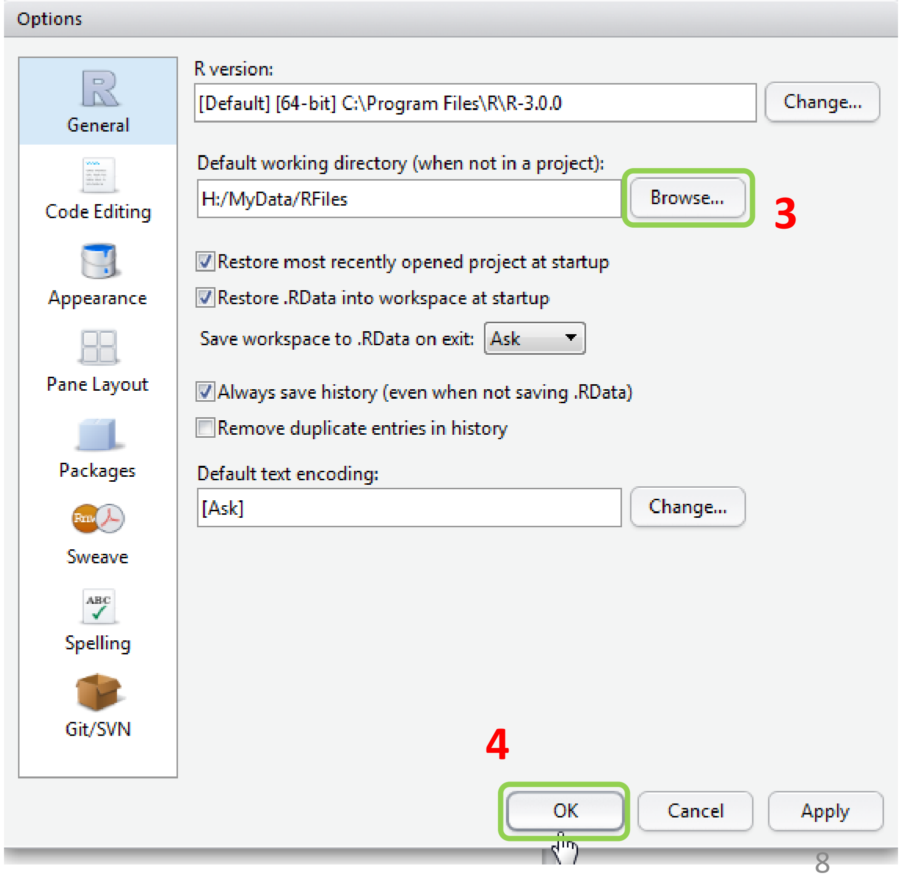


1

2

Console H:/MyData/RFiles/

```
R version 3.0.0 (2013-04-03) -- "Masked Ma
Copyright (C) 2013 The R Foundation for St
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTE
You are welcome to redistribute it under c
Type 'license()' or 'licence()' for distri
```



R version: [Default] [64-bit] C:\Program Files\R\R-3.0.0

Default working directory (when not in a project): H:/MyData/RFiles

3

☑ Restore most recently opened project at startup

☑ Restore .RData into workspace at startup

Save workspace to .RData on exit: Ask

☑ Always save history (even when not saving .RData)

□ Remove duplicate entries in history

Default text encoding: [Ask]

4

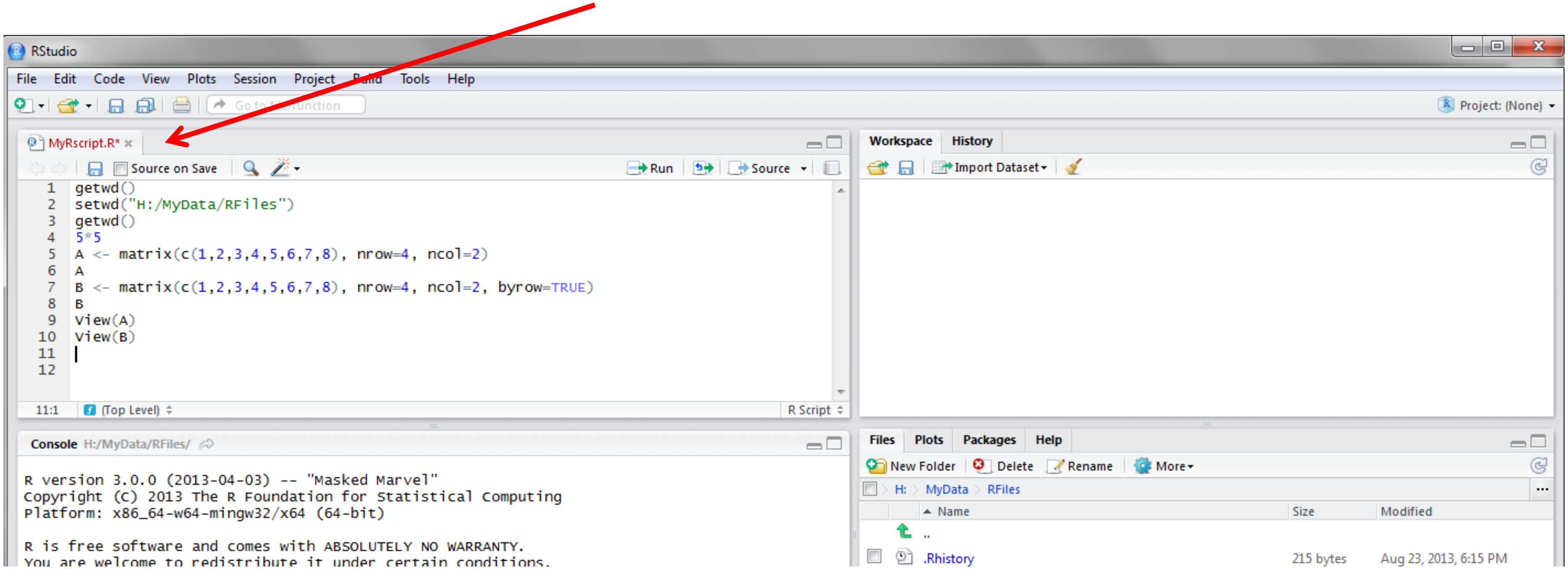
Every time you open RStudio, it goes to a default directory. You can change the default to a folder where you have your datafiles so you do not have to do it every time. In the menu go to Tools->Options

R script (1)

The usual Rstudio screen has four windows:

1. Console.
2. Workspace and history.
3. Files, plots, packages and help.
4. The R script(s) and data view.

The R script is where you keep a record of your work. For Stata users this would be like the do-file, for SPSS users is like the syntax and for SAS users the SAS program.



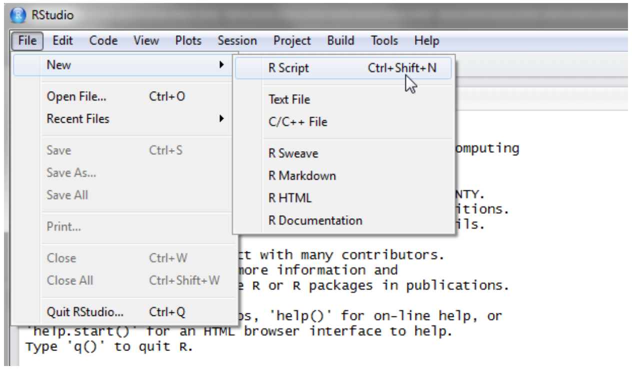
```
getwd()
setwd("H:/MyData/RFiles")
getwd()
5*5
A <- matrix(c(1,2,3,4,5,6,7,8), nrow=4, ncol=2)
A
B <- matrix(c(1,2,3,4,5,6,7,8), nrow=4, ncol=2, byrow=TRUE)
B
View(A)
View(B)
```

```
R version 3.0.0 (2013-04-03) -- "Masked Marvel"
Copyright (C) 2013 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
```

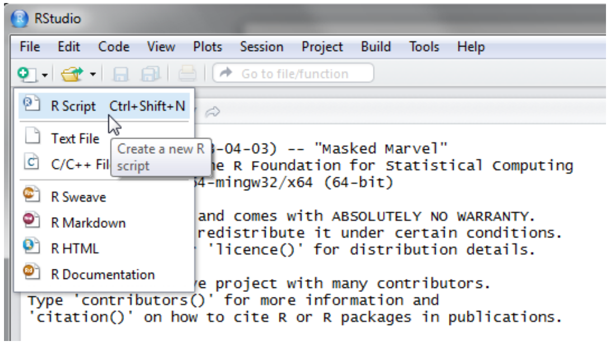
Name	Size	Modified
..		
.Rhistory	215 bytes	Aug 23, 2013, 6:15 PM

R script (2)

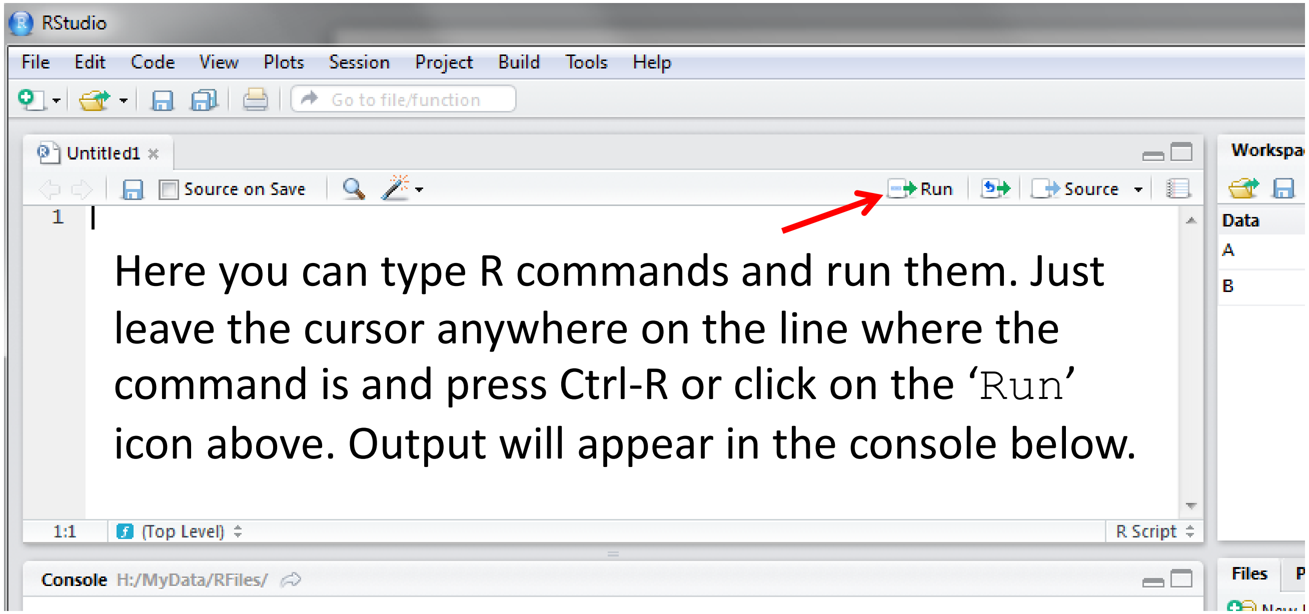
To create a new R script you can either go to `File -> New -> R Script`, or click on the icon with the "+" sign and select "R Script", or simply press Ctrl+Shift+N. Make sure to save the script.



```
omputing
NTY.
itions.
ils.
t with many contributors.
more information and
e R or R packages in publications.
s, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
```



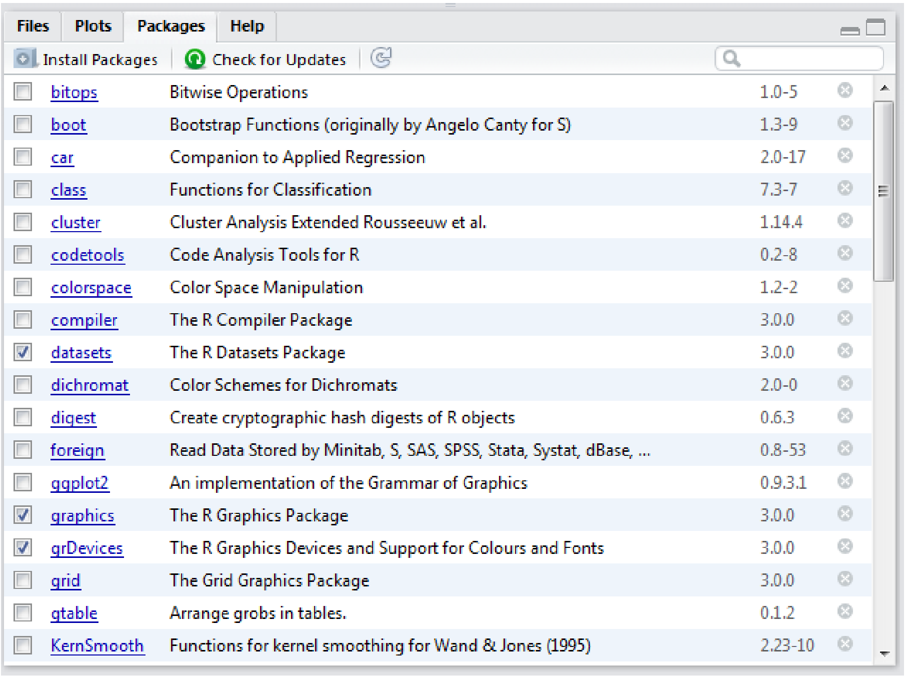
```
-04-03) -- "Masked Marvel"
e R Foundation for Statistical Computing
4-mingw32/x64 (64-bit)
and comes with ABSOLUTELY NO WARRANTY.
redistribute it under certain conditions.
'licence()' for distribution details.
e project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
```



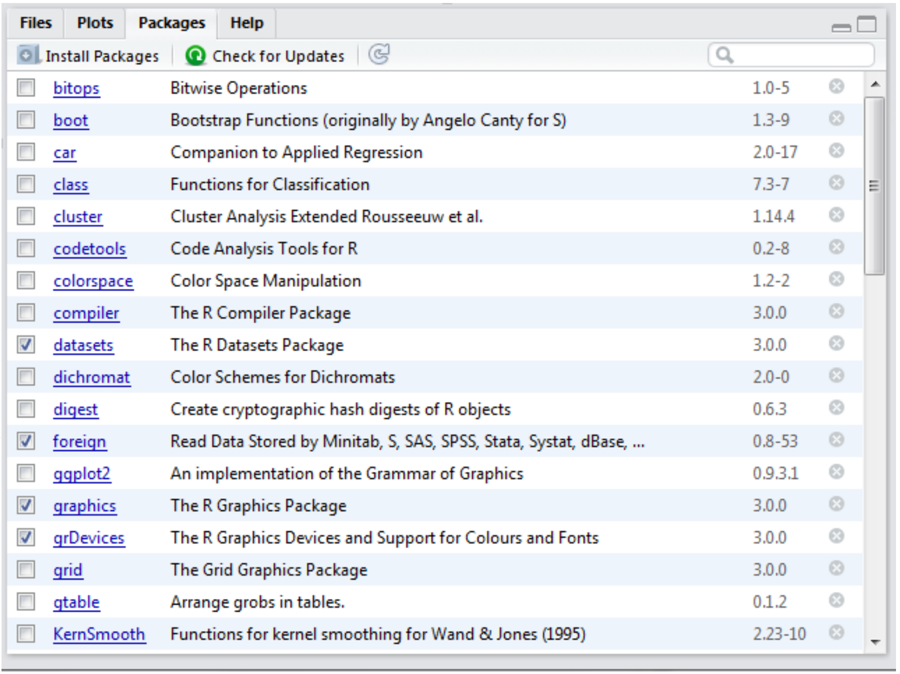
Here you can type R commands and run them. Just leave the cursor anywhere on the line where the command is and press Ctrl-R or click on the 'Run' icon above. Output will appear in the console below.

Packages tab

The package tab shows the list of add-ons included in the installation of RStudio. If checked, the package is loaded into R, if not, any command related to that package won’t work, you will need select it. You can also install other add-ons by clicking on the ‘Install Packages’ icon. Another way to activate a package is by typing, for example, `library(foreign)`. This will automatically check the `--foreign` package (it helps bring data from proprietary formats like Stata, SAS or SPSS).

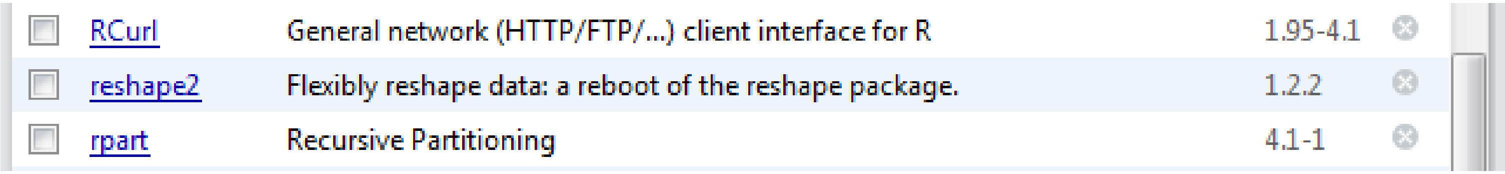


□	bitops	Bitwise Operations	1.0-5
□	boot	Bootstrap Functions (originally by Angelo Canty for S)	1.3-9
□	car	Companion to Applied Regression	2.0-17
□	class	Functions for Classification	7.3-7
□	cluster	Cluster Analysis Extended Rousseeuw et al.	1.14.4
□	codetools	Code Analysis Tools for R	0.2-8
□	colorspace	Color Space Manipulation	1.2-2
□	compiler	The R Compiler Package	3.0.0
☑	datasets	The R Datasets Package	3.0.0
□	dichromat	Color Schemes for Dichromats	2.0-0
□	digest	Create cryptographic hash digests of R objects	0.6.3
□	foreign	Read Data Stored by Minitab, S, SAS, SPSS, Stata, Systat, dBase, ...	0.8-53
□	ggplot2	An implementation of the Grammar of Graphics	0.9.3.1
☑	graphics	The R Graphics Package	3.0.0
☑	grDevices	The R Graphics Devices and Support for Colours and Fonts	3.0.0
□	grid	The Grid Graphics Package	3.0.0
□	gtable	Arrange grobs in tables.	0.1.2
□	KernSmooth	Functions for kernel smoothing for Wand & Jones (1995)	2.23-10



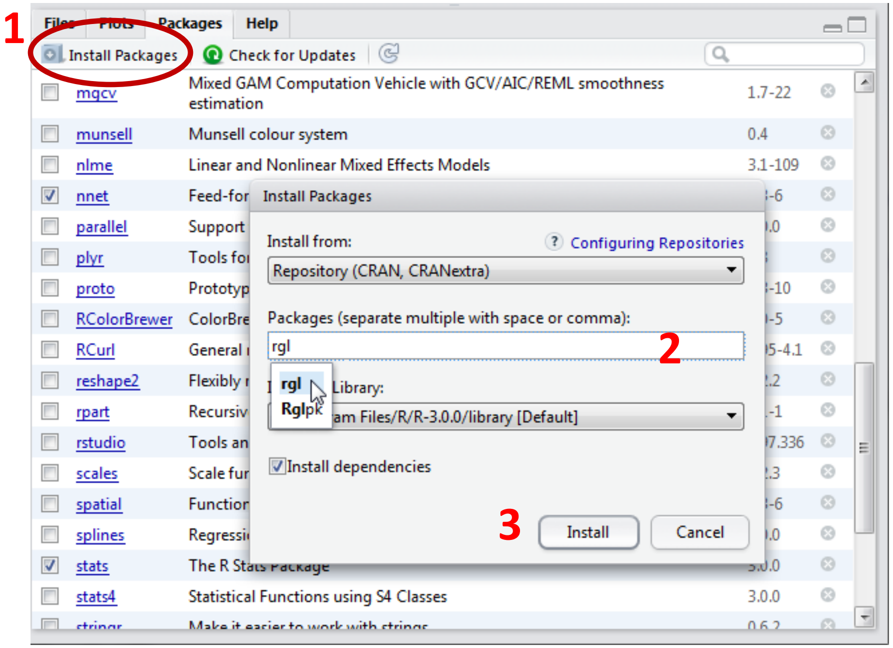
□	bitops	Bitwise Operations	1.0-5
□	boot	Bootstrap Functions (originally by Angelo Canty for S)	1.3-9
□	car	Companion to Applied Regression	2.0-17
□	class	Functions for Classification	7.3-7
□	cluster	Cluster Analysis Extended Rousseeuw et al.	1.14.4
□	codetools	Code Analysis Tools for R	0.2-8
□	colorspace	Color Space Manipulation	1.2-2
□	compiler	The R Compiler Package	3.0.0
☑	datasets	The R Datasets Package	3.0.0
□	dichromat	Color Schemes for Dichromats	2.0-0
□	digest	Create cryptographic hash digests of R objects	0.6.3
☑	foreign	Read Data Stored by Minitab, S, SAS, SPSS, Stata, Systat, dBase, ...	0.8-53
□	ggplot2	An implementation of the Grammar of Graphics	0.9.3.1
☑	graphics	The R Graphics Package	3.0.0
☑	grDevices	The R Graphics Devices and Support for Colours and Fonts	3.0.0
□	grid	The Grid Graphics Package	3.0.0
□	gtable	Arrange grobs in tables.	0.1.2
□	KernSmooth	Functions for kernel smoothing for Wand & Jones (1995)	2.23-10

Installing a package



□	RCurl	General network (HTTP/FTP/...) client interface for R	1.95-4.1
□	reshape2	Flexibly reshape data: a reboot of the reshape package.	1.2.2
□	rpart	Recursive Partitioning	4.1-1

Before



□	mgcv	Mixed GAM Computation Vehicle with GCV/AIC/REML smoothness estimation	1.7-22
□	munsell	Munsell colour system	0.4
□	nlme	Linear and Nonlinear Mixed Effects Models	3.1-109
☑	nnet	Feed-for	-6
□	parallel	Support	.0
□	plyr	Tools fo	
□	proto	Prototyp	-10
□	RColorBrewer	ColorBre	-5
□	RCurl	General	5-4.1
□	reshape2	Flexibly	.2
□	rpart	Recursiv	-1
□	rstudio	Tools an	7.336
□	scales	Scale fu	.3
□	spatial	Function	-6
□	splines	Regressi	.0
☑	stats	The R Stats Package	3.0.0
□	stats4	Statistical Functions using S4 Classes	3.0.0
□	stringr	Make it easier to work with strings	0.6.2

Install Packages

Install from: Repository (CRAN, CRANextra) ? Configuring Repositories

Packages (separate multiple with space or comma): rgl

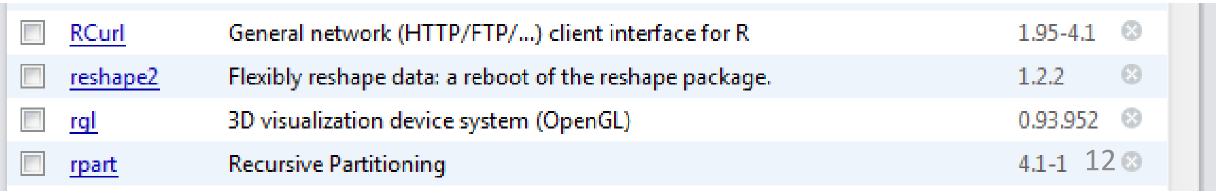
rgl
Rglpk

Install to Library: ...am Files/R/R-3.0.0/library [Default]

☑ Install dependencies

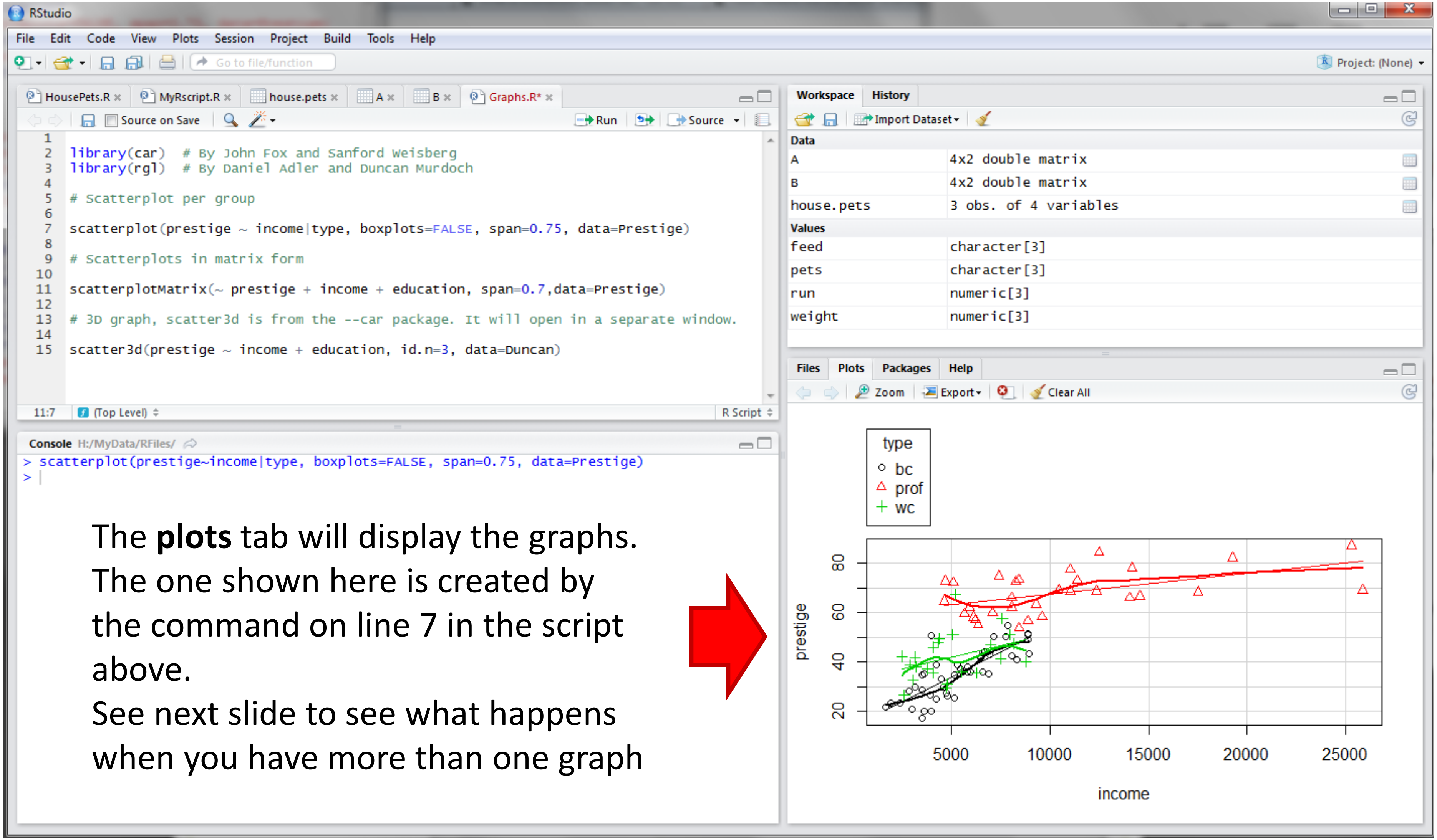
We are going to install the package – `rgl` (useful to plot 3D images). It does not come with the original R install.
Click on "`Install Packages`", write the name in the pop-up window and click on "`Install`".

After



□	RCurl	General network (HTTP/FTP/...) client interface for R	1.95-4.1
□	reshape2	Flexibly reshape data: a reboot of the reshape package.	1.2.2
□	rgl	3D visualization device system (OpenGL)	0.93.952
□	rpart	Recursive Partitioning	4.1-1

Plots tab (1)


RStudio
File Edit Code View Plots Session Project Build Tools Help
Go to file/function
Project: (None)
HousePets.R MyRscript.R house.pets A B Graphs.R*
Source on Save
Run
Source
Workspace History
Import Dataset
Files Plots Packages Help
Zoom Export Clear All
11:7 (Top Level) R Script
Console H:/MyData/RFiles/
type
bc
prof
wc
prestige
income
20
40
60
80
5000
10000
15000
20000
25000

```
1
2   library(car)  # By John Fox and Sanford Weisberg
3   library(rgl)  # By Daniel Adler and Duncan Murdoch
4
5   # Scatterplot per group
6
7   scatterplot(prestige ~ income|type, boxplots=FALSE, span=0.75, data=Prestige)
8
9   # Scatterplots in matrix form
10
11  scatterplotMatrix(~ prestige + income + education, span=0.7,data=Prestige)
12
13  # 3D graph, scatter3d is from the --car package. It will open in a separate window.
14
15  scatter3d(prestige ~ income + education, id.n=3, data=Duncan)
```

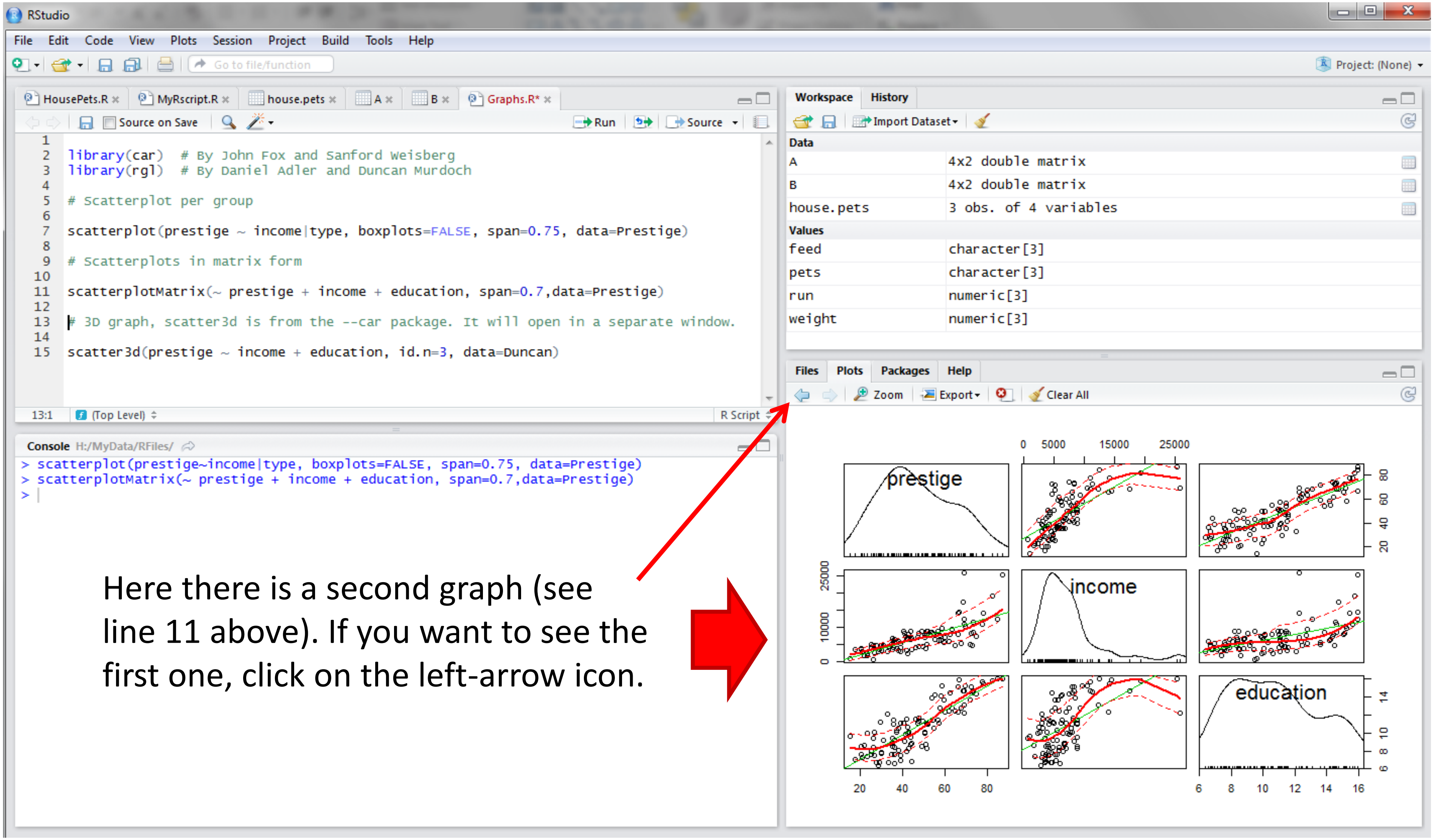
Data	
A	4x2 double matrix
B	4x2 double matrix
house.pets	3 obs. of 4 variables
Values	
feed	character[3]
pets	character[3]
run	numeric[3]
weight	numeric[3]

```
> scatterplot(prestige~income|type, boxplots=FALSE, span=0.75, data=Prestige)
> |
```

The **plots** tab will display the graphs. The one shown here is created by the command on line 7 in the script above.

See next slide to see what happens when you have more than one graph

Plots tab (2)



```
1
2  library(car)  # By John Fox and Sanford Weisberg
3  library(rgl)  # By Daniel Adler and Duncan Murdoch
4
5  # Scatterplot per group
6
7  scatterplot(prestige ~ income|type, boxplots=FALSE, span=0.75, data=Prestige)
8
9  # Scatterplots in matrix form
10
11 scatterplotMatrix(~ prestige + income + education, span=0.7,data=Prestige)
12
13 # 3D graph, scatter3d is from the --car package. It will open in a separate window.
14
15 scatter3d(prestige ~ income + education, id.n=3, data=Duncan)
```

Data	
A	4x2 double matrix
B	4x2 double matrix
house.pets	3 obs. of 4 variables
Values	
feed	character[3]
pets	character[3]
run	numeric[3]
weight	numeric[3]

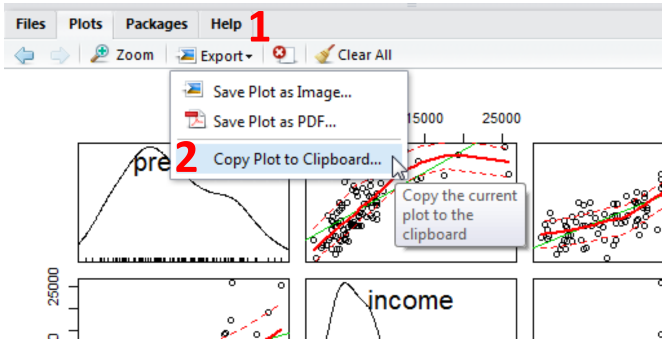
Console H:/MyData/RFiles/

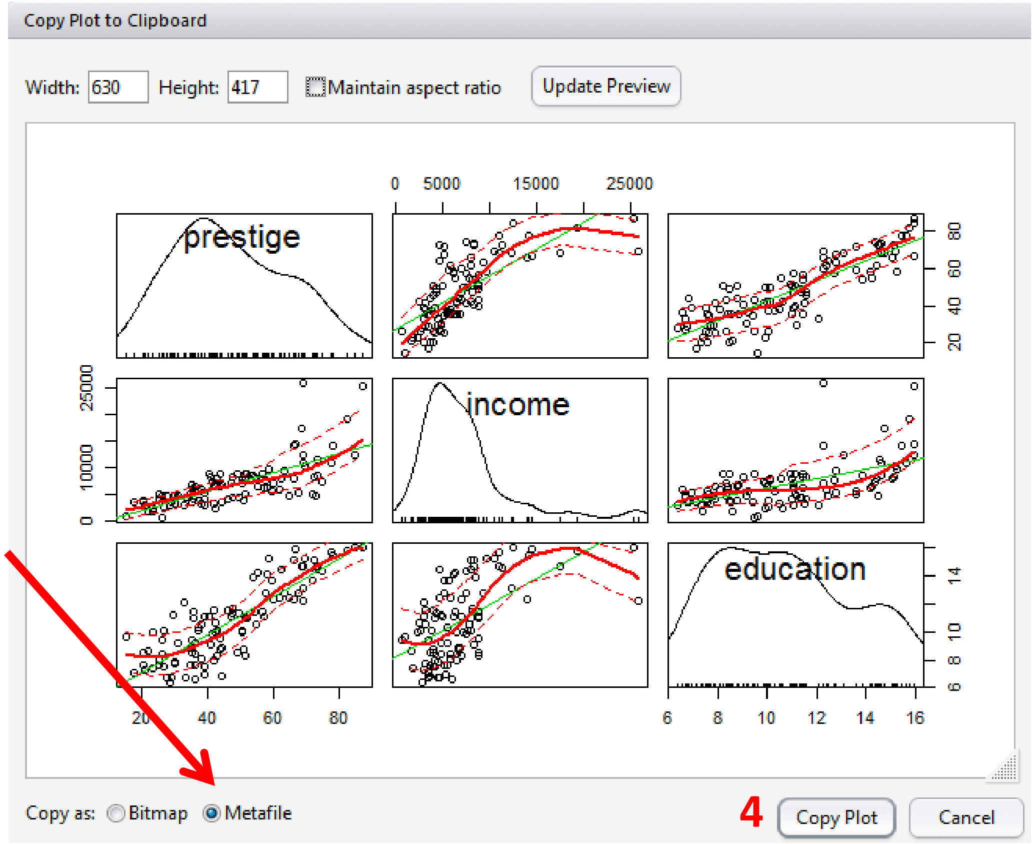
```
> scatterplot(prestige~income|type, boxplots=FALSE, span=0.75, data=Prestige)
> scatterplotMatrix(~ prestige + income + education, span=0.7,data=Prestige)
> |
```

Here there is a second graph (see line 11 above). If you want to see the first one, click on the left-arrow icon.

Plots tab (3) – Graphs export

To extract the graph, click on “`Export`” where you can save the file as an image (PNG, JPG, etc.) or as PDF, these options are useful when you only want to share the graph or use it in a LaTeX document. Probably, the easiest way to export a graph is by copying it to the clipboard and then paste it directly into your Word document.





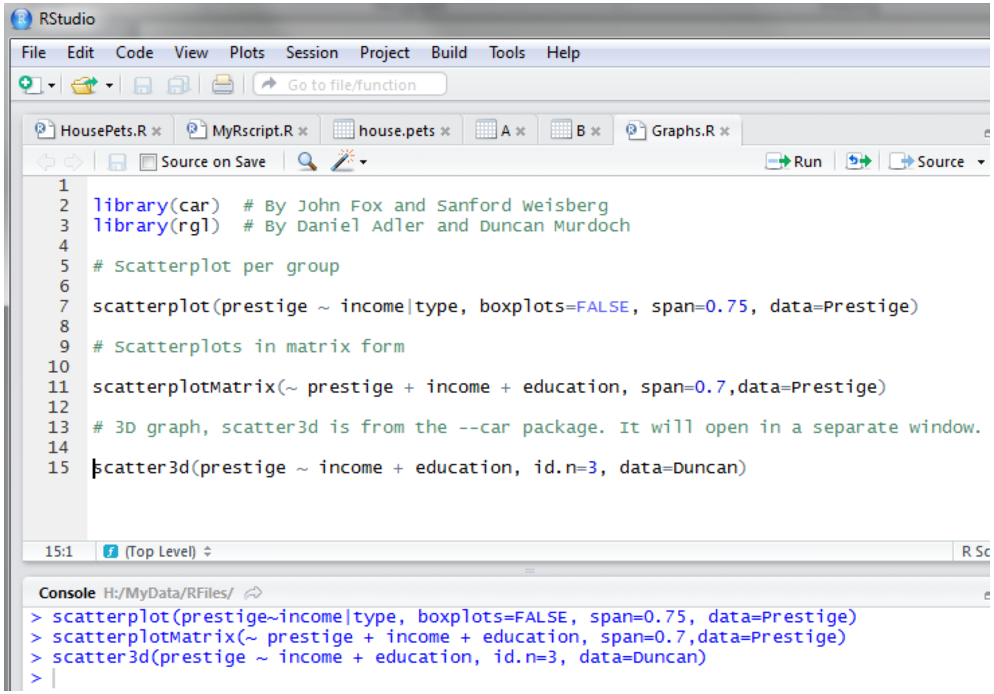
Width: 630 Height: 417 □ Maintain aspect ratio

Copy as: ○ Bitmap ● Metafile

3 Make sure to select ‘`Metafile`’

5 Paste it into your Word document

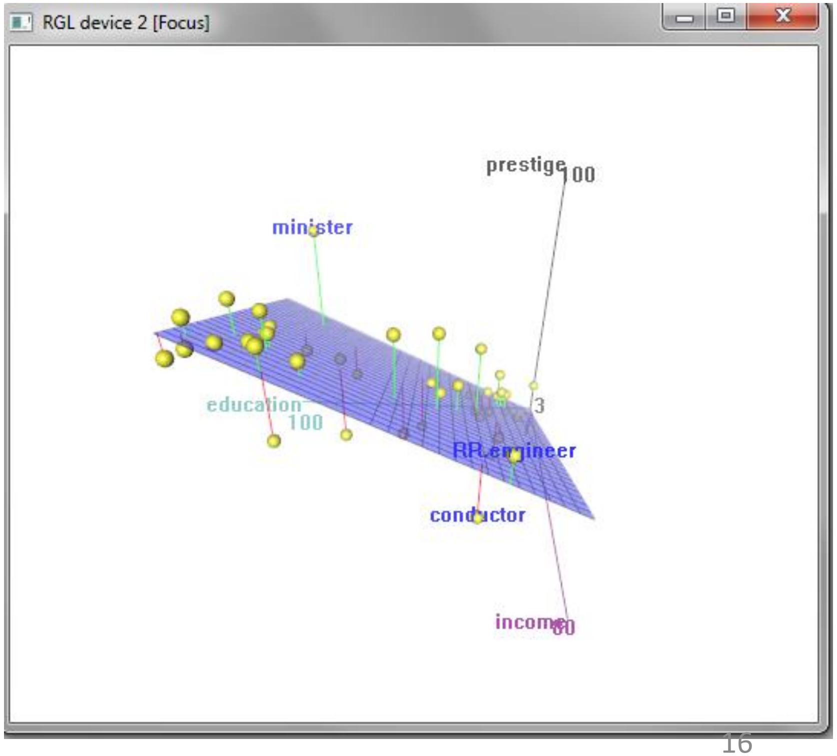
3D graphs



```
 1
 2  library(car)  # By John Fox and Sanford Weisberg
 3  library(rgl)  # By Daniel Adler and Duncan Murdoch
 4
 5  # Scatterplot per group
 6
 7  scatterplot(prestige ~ income|type, boxplots=FALSE, span=0.75, data=Prestige)
 8
 9  # Scatterplots in matrix form
10
11  scatterplotMatrix(~ prestige + income + education, span=0.7,data=Prestige)
12
13  # 3D graph, scatter3d is from the --car package. It will open in a separate window.
14
15  scatter3d(prestige ~ income + education, id.n=3, data=Duncan)
```

```
> scatterplot(prestige~income|type, boxplots=FALSE, span=0.75, data=Prestige)
> scatterplotMatrix(~ prestige + income + education, span=0.7,data=Prestige)
> scatter3d(prestige ~ income + education, id.n=3, data=Duncan)
> 
```

3D graphs will display on a separate screen (see line 15 above). You won’t be able to save it, but after moving it around, once you find the angle you want, you can screenshot it and paste it to you Word document.


RGL device 2 [Focus]
prestige
minister
education
RR.engineer
conductor
income