



Introduction to RStudio

(v 3.6)

Oscar Torres-Reyna

otorres@princeton.edu

Installing RStudio

1. Download and install R, see here:
<https://cran.r-project.org/>
2. Download and install RStudio Desktop (free version), see here:
<https://posit.co/download/rstudio-desktop/>

If RStudio does not recognize the latest installed version of R, see here:

<https://support.posit.co/hc/en-us/articles/200486138-Using-Different-Versions-of-R>

RStudio screen (first time)

The screenshot shows the RStudio interface with the following components:

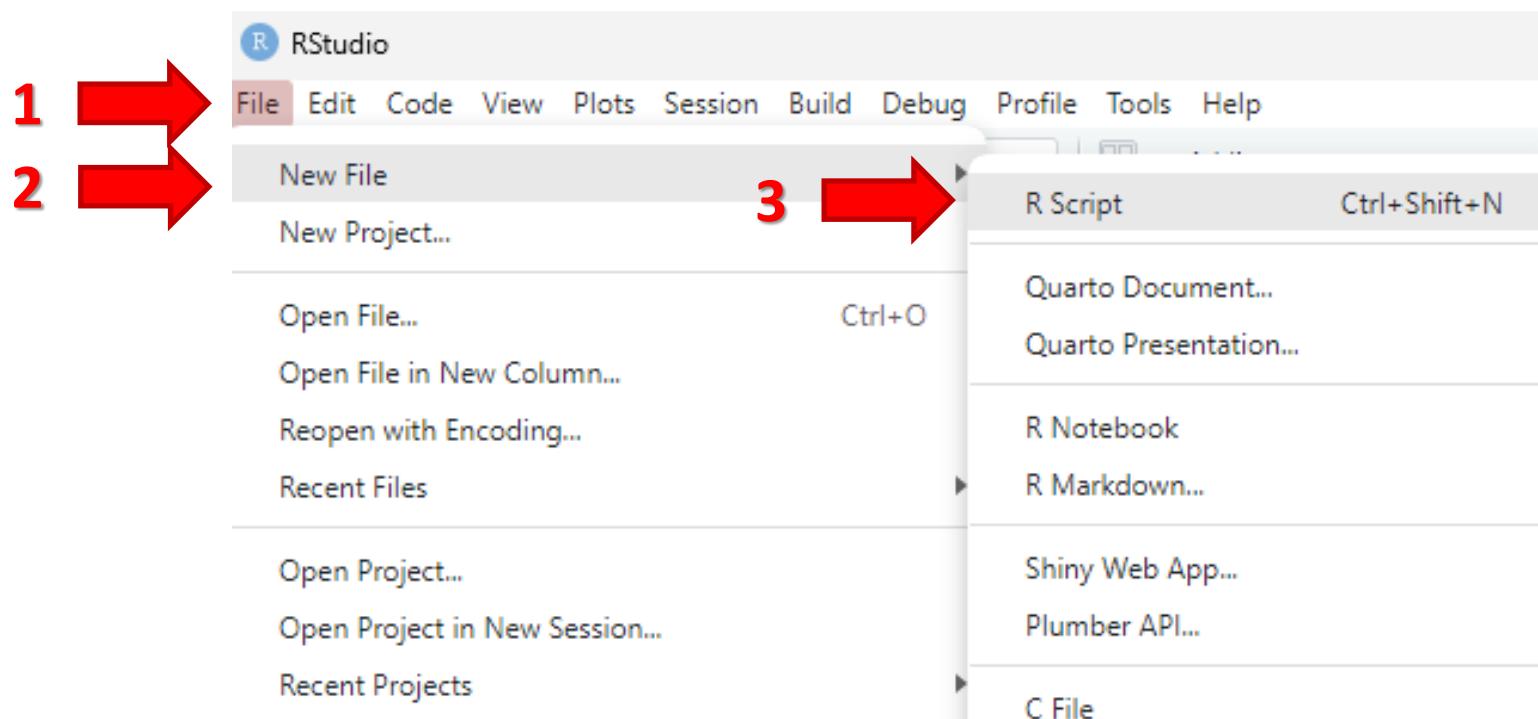
- Source** tab: Shows a blank code editor area.
- Console** tab: Shows the R startup message and a command-line interaction where `> 3 + 6` results in `[1] 9`.
- Environment** tab: Shows a message "Environment is empty".
- Files** tab: Shows a file browser with a folder named "Global Environment".

The console is where you can type commands and see output

The **environment** tab shows all the active objects (see next slide). The **history** tab shows a list of commands used in the session

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 list the packages available, can install packages.

Open an R Script window



[See next slide]

RStudio screen (w/R Script)

The screenshot shows the RStudio interface. At the top is a menu bar with File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu is a toolbar with icons for file operations like Open, Save, and Print, along with Go to file/function and Addins. The main workspace is titled Untitled1. A red arrow points from the text above to the 'Run' icon in the toolbar. To the right is the Environment tab, which displays the Global Environment and notes that it is empty. Below the workspace is a navigation bar with tabs for Files, Plots, Packages, Help, Viewer, and Presentation. The bottom left shows the R console output, which includes the R version information, copyright notice, and various system and R-specific messages. The bottom right contains descriptive text about the different tabs.

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.

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

The **environment** tab shows all the active objects (see next slide). The **history** tab shows a list of commands used in the session

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 lists the packages available, can install packages.

Set working directory

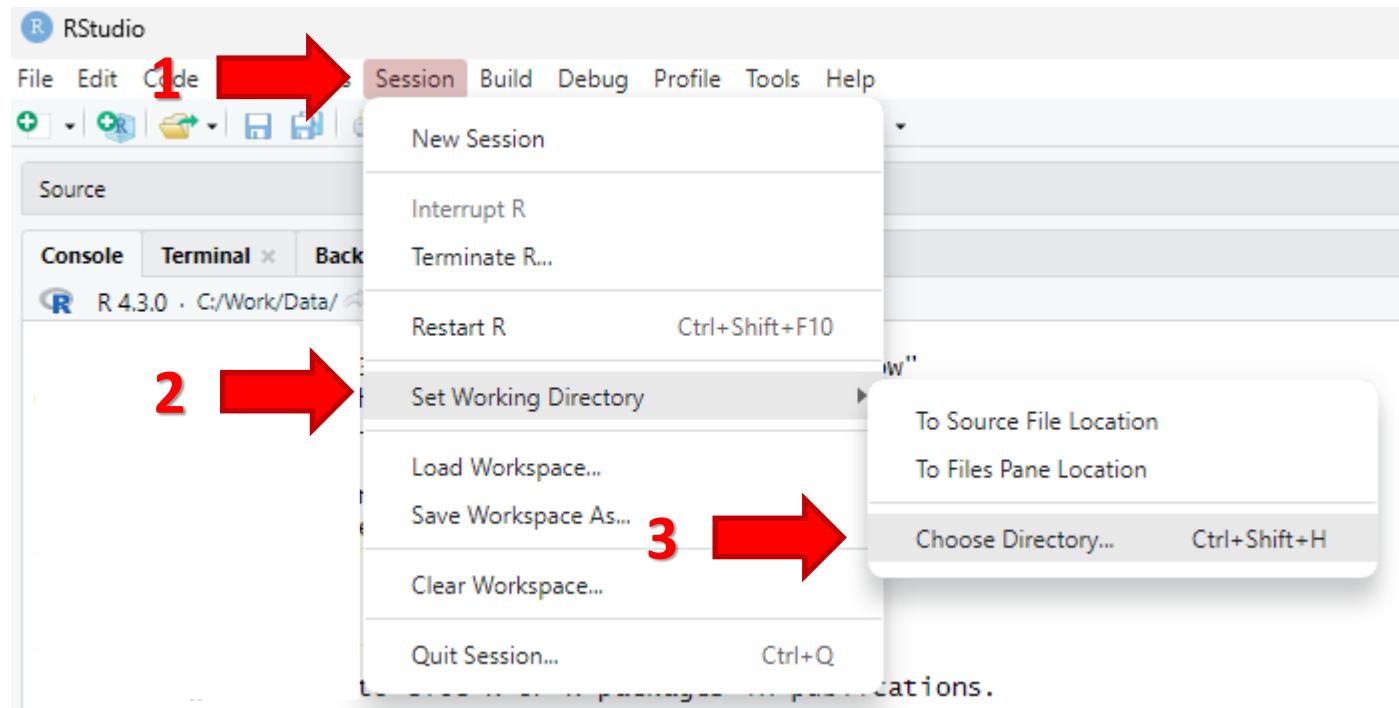
The working directory is the folder where your files are saved and where RStudio will export files. You can check the current working directory by typing the following:

```
getwd()
```

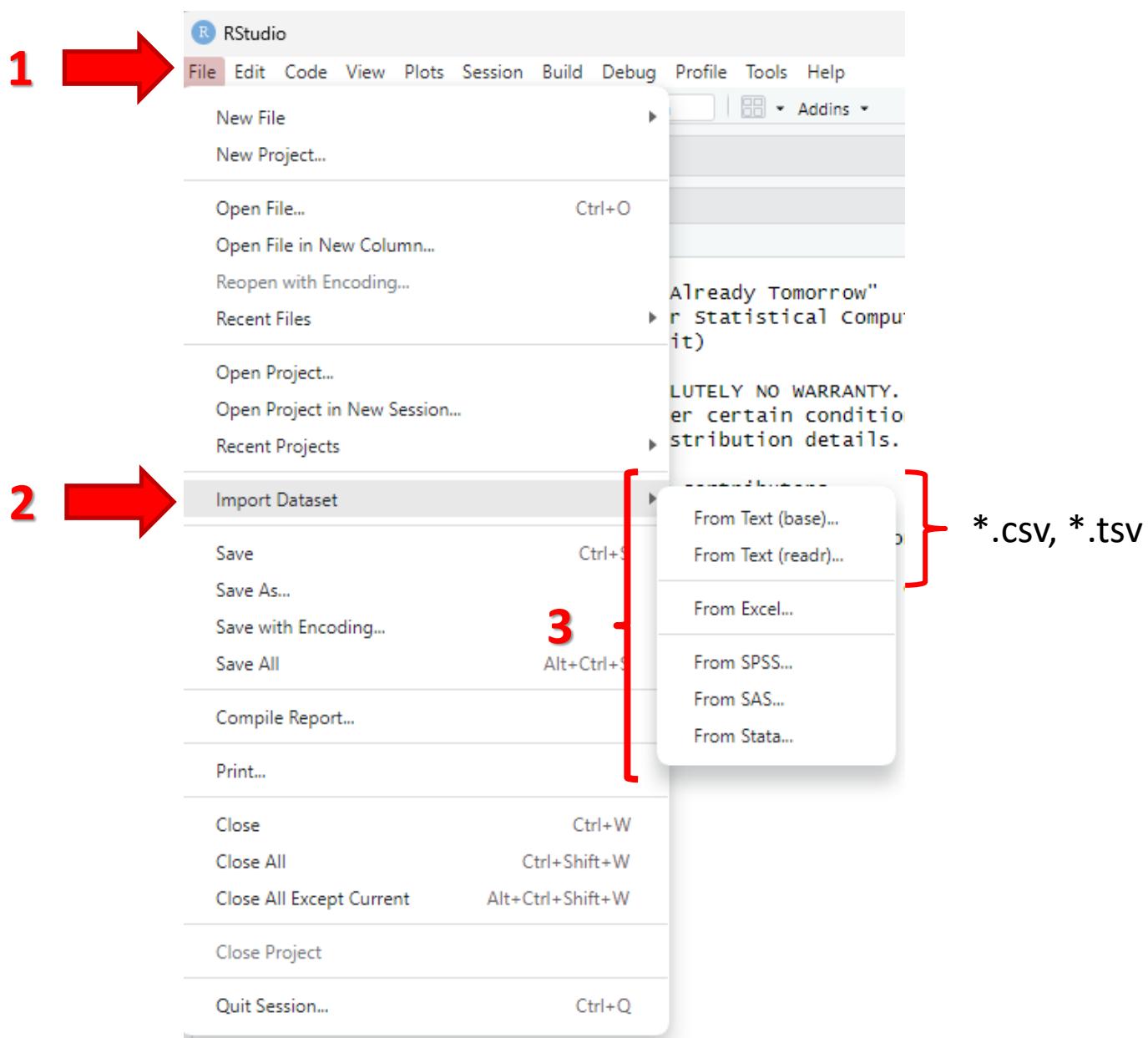
You can change the working directory by typing (if using Windows):

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

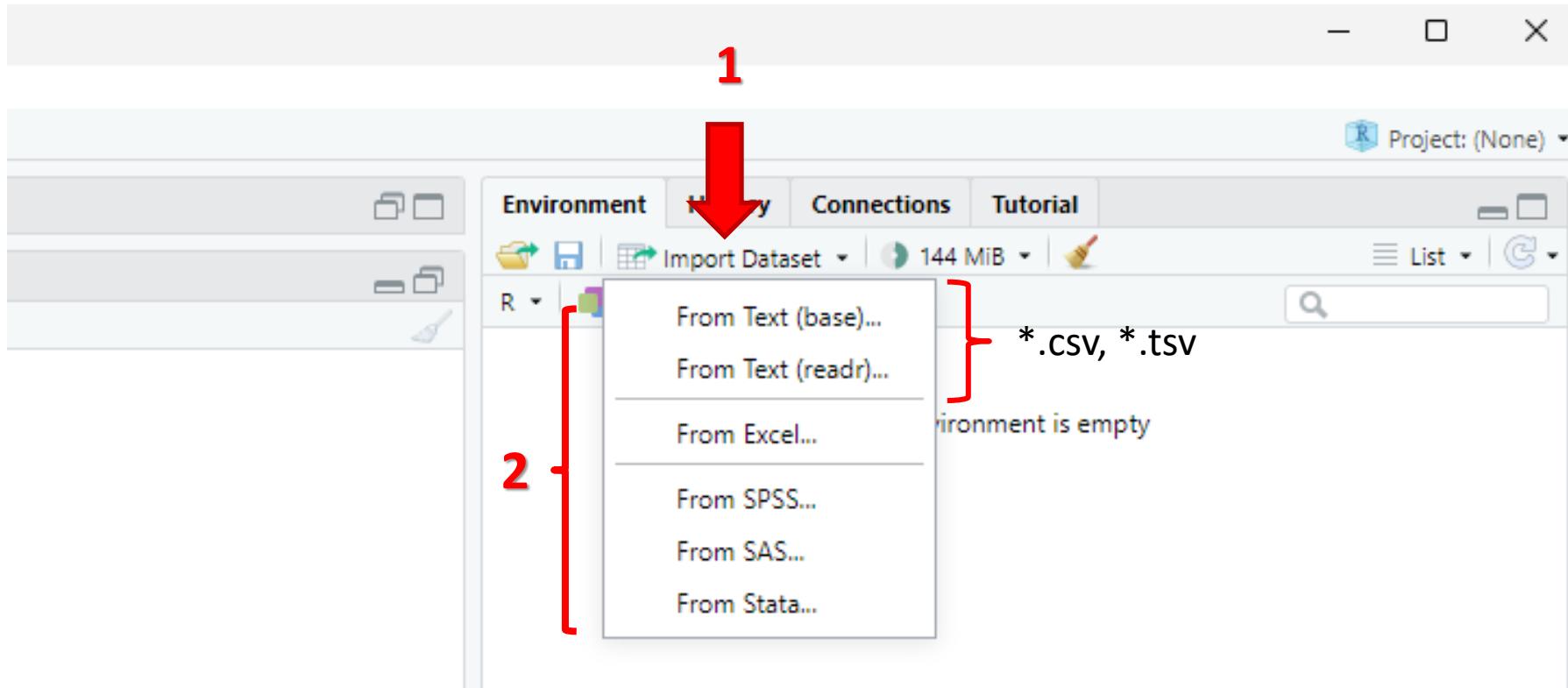
Or you can use the menu:



Import files into RStudio (1)

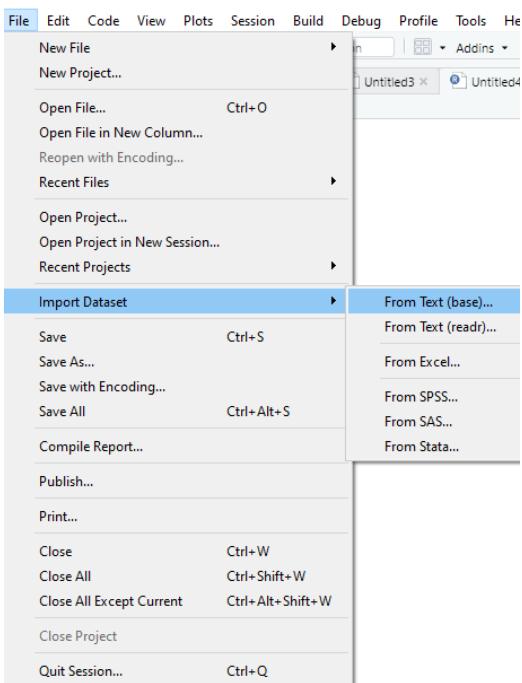


Import files into RStudio (2)



NOTE: When importing files from Excel or ASCII format (*.csv, *.tsv) make sure the data is in the format shown in slides 2 and/or 3 in this document:

<https://www.princeton.edu/~otorres/DataPrep101.pdf>



After loading the data, change name to something short here

If your data has column titles, check 'Yes' here

Import Dataset

Name:

Encoding:

Heading: Yes No

Row names:

Separator:

Decimal:

Quote:

Comment:

na.strings:

Strings as factors

Input File

```
Country,x1995,x1996,x1997,x1998,x1999,x2000,x2001,x2002,x2003
A,,8000.01,8212.90,7847.36,7702.89,7288.48,6430.98,6932.45,7
B,18268.01,18738.99,19360.46,20151.42,20715.54,20866.90,21364
C,21088.14,21608.14,21988.64,22739.28,23436.61,24194.85,24300
D,313.74,321.36,331.76,342.12,351.70,365.33,377.15,386.26,398
E,21123.66,21659.55,22299.13,22972.31,23613.87,24150.86,24788
F,29941.64,30703.73,31716.04,32671.27,33748.21,34599.47,34483
G,4891.60,5063.81,5328.88,5512.59,5647.06,5934.98,5864.12,585
```

Data Frame

Country	x1995	x1996	x1997	x1998	x1999
A	NA	NA	8000.01	8212.90	7847.36
B	18268.01	18738.99	19360.46	20151.42	20715.54
C	21088.14	21608.14	21988.64	22739.28	23436.61
D	313.74	321.36	331.76	342.12	351.70
E	21123.66	21659.55	22299.13	22972.31	23613.87
F	29941.64	30703.73	31716.04	32671.27	33748.21
G	4891.60	5063.81	5328.88	5512.59	5647.06

After importing,
copy the code to
the R script

Click here to import the data

Import

Cancel

Using import data From Text (base)

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

New File
New Project...
Open File... Ctrl+O
Open File in New Column...
Reopen with Encoding...
Recent Files
Open Project...
Open Project in New Session...
Recent Projects

Import Dataset
Save Ctrl+S
Save As...
Save with Encoding...
Save All Ctrl+Alt+S
Compile Report...
Publish...
Print...
Close Ctrl+W
Close All Ctrl+Shift+W
Close All Except Current Ctrl+Alt+Shift+W
Close Project
Quit Session... Ctrl+Q

From Text (base)...
From Text (readr)...
From Excel...
From SPSS...
From SAS...
From Stata...

Import Text Data
File/URL: ~/GDP_OneVar.csv

Data Preview:

Country (character)	x1995 (double)	x1996 (double)	x1997 (double)	x1998 (double)	x1999 (double)	x2000 (double)	x2001 (double)	x2002 (double)	x2003 (double)	x2004 (double)	x2005 (double)
A	NA	NA	8000.01	8212.90	7847.36	7702.89	7288.48	6430.98	6932.45	7486.24	8094.17
B	18268.01	18738.99	19360.46	20151.42	20715.54	20866.90	21364.02	21801.41	22404.59	22676.26	23039.43
C	21068.14	21608.14	21988.64	22739.28	23436.61	24194.85	24300.57	24411.48	24650.02	25076.01	25346.01
D	313.74	321.36	331.76	342.12	351.70	365.33	377.15	386.26	398.66	415.96	432.63
E	21123.66	21659.55	22299.13	22972.31	23613.87	24150.86	24788.69	25368.87	25885.48	26582.19	26890.73
F	29941.64	30703.73	31716.04	32671.27	33748.21	34599.47	34483.98	34669.47	35312.75	36450.55	37267.33
G	4891.60	5063.81	5328.88	5512.59	5647.06	5934.98	5864.12	5852.99	5872.29	6055.92	6162.84

Change the name to something short here

Previewing first 50 entries.

Import Options:

Name: mydata	First Row as Names	Delimiter: Comma	Escape: None
Skip: 0	Trim Spaces	Quotes: Default	Comment: Default
Open Data Viewer		Locale: Configure...	NA: Default

Reading rectangular data using readr

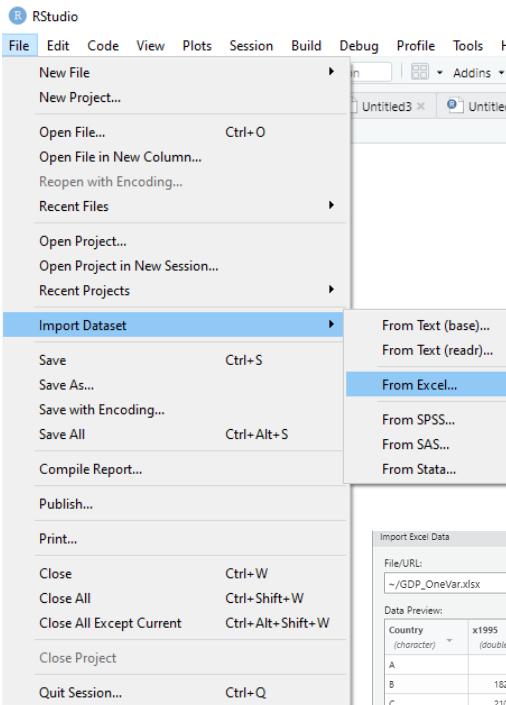
Code Preview:

```
library(readr)
mydata <- read_csv("~/GDP_OneVar.csv")
View(mydata)
```

Click here to import the data

Click here to find the file

Import Cancel



Using import data From Excel

Click here to find the file

If the file has more than one sheet, select the one you need here...

Change the name to something short here

R code shown here (copy it to the R script after importing)

Click here to import the data

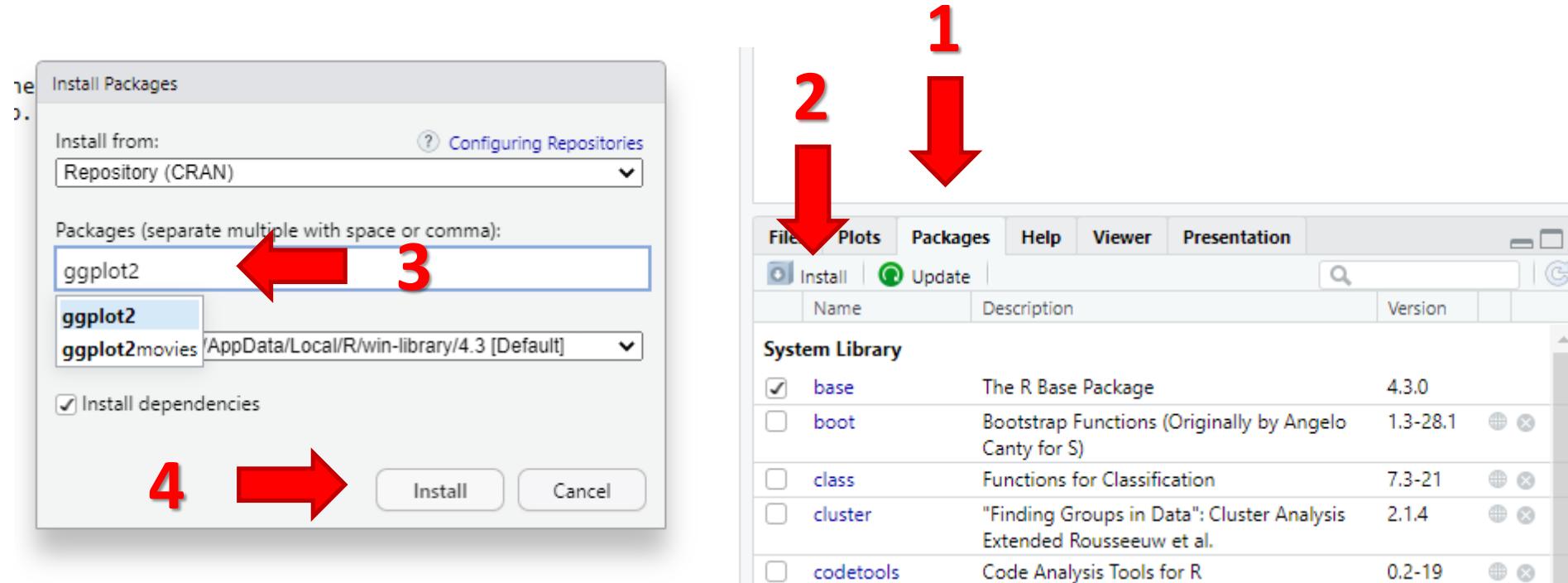
Import Options

Name: mydata Max Rows: First Row as Names:
Sheet: Default Skip: 0 Open Data Viewer:
Range: A1:D10 NA:

Code Preview:

```
library(readxl)
mydata <- read_excel("~/GDP_OneVar.xlsx")
View(mydata)
```

Installing a package



Once a package is installed, no need to install it again until a new version of R is installed.
To activate a package type

```
library(name_of_package)
```

Data visualization

R RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Untitled1 x Graphs.R* x Go to file/function Addins

1 library(car) # By John Fox and Sanford Weisberg
2 library(rgl) # By Daniel Adler and Duncan Murdoch
3
4 # scatterplot
5
6 ● 7 scatterplot(prestige ~ income, data=Prestige) ← 1 Run
8
9 # scatterplot per group
10
11 scatterplot(prestige ~ income|type, data=Prestige)
12
13 # Scatterplots in matrix form
14
15 scatterplotMatrix(~ prestige + income + education, span=0.7, data=Prestige)
16
17 # 3D graph, scatter3d is from the --car package. It will open in the Viewer window
18
19 scatter3d(prestige ~ income + education, id.n=3, data=Duncan)
20 rglwidget()

7:1 (Top Level) ⇣ R Script ⇣

Console Terminal x Background Jobs x

R 4.3.0 - C:/Work/Data/ ↵
> scatterplot(prestige ~ income, data=Prestige)
> |

[See next slide]

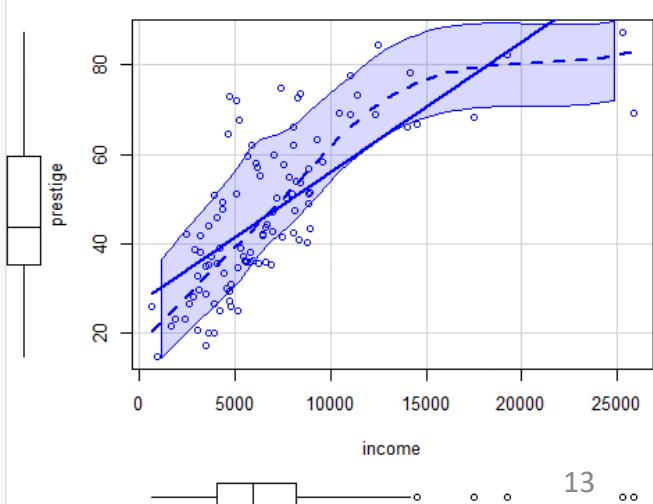
Environment History Connections Tutorial

Import Dataset 333 MB Global Environment

Environment is empty

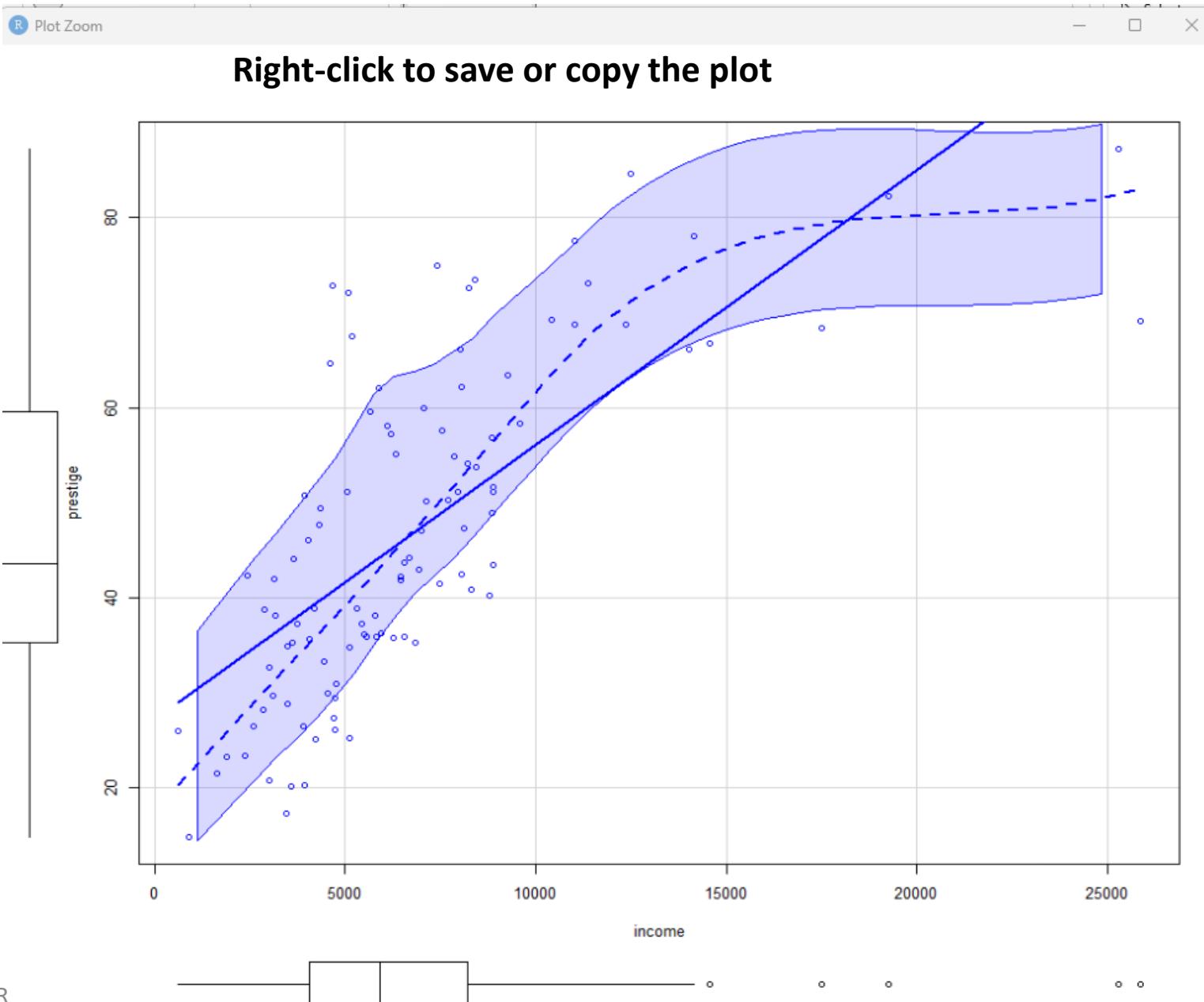
Files Plots Packages Help Viewer Presentation

Zoom Export



13

Data visualization



Data visualization

R RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Untitled1 x Graphs.R x Go to file/function Addins

```
1 library(car) # By John Fox and Sanford Weisberg
2 library(rgl) # By Daniel Adler and Duncan Murdoch
3
4 # Scatterplot
5
6 scatterplot(prestige ~ income, data=Prestige)
7 # Scatterplot per group
8
9 scatterplot(prestige ~ income|type, data=Prestige)
10
11 # Scatterplots in matrix form
12
13 scatterplotMatrix(~ prestige + income + education, span=0.7, data=Prestige)
14
15 # 3D graph, scatter3d is from the --car package. It will open in the Viewer window
16
17 scatter3d(prestige ~ income + education, id.n=3, data=Duncan)
18 rglwidget()
```

● 7:1 (Top Level) ◊

Console Terminal Background Jobs

R 4.3.0 - C:/Work/Data/ > scatterplot(prestige ~ income, data=Prestige)

[See next slide]

1 Run

Environment History Connections Tutorial

Import Dataset 333 MB Global Environment

Environment is empty

2 click on Export to save or copy the plot.

Files Plots Packages Export Viewer Presentation

Zoom Export

The scatterplot displays the relationship between income (X-axis, ranging from 0 to 25,000) and prestige (Y-axis, ranging from 0 to 80). The data points show a positive correlation. A solid blue line represents the linear regression fit, and a light blue shaded area indicates the confidence interval. A dashed blue line is also visible. A red arrow points to the 'Run' button in the top left of the RStudio interface, with the text '1 Run' overlaid. Another red arrow points to the 'Export' button in the bottom navigation bar, with the text '2 click on Export to save or copy the plot.' overlaid. The bottom right corner shows a small portion of a box plot for the 'prestige' variable.

OTR 15

Data visualization (3D graphics)

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Untitled1 x Graphs.R* x Go to file/function Addins

```
1 library(car) # By John Fox and Sanford Weisberg
2 library(rgl) # By Daniel Adler and Duncan Murdoch
3
4 # scatterplot
5
6 scatterplot(prestige ~ income, data=Prestige)
7
8 # scatterplot per group
9
10 scatterplot(prestige ~ income|type, data=Prestige)
11
12 # scatterplots in matrix form
13
14 scatterplotMatrix(~ prestige + income + education, span=0.7, data=Prestige)
15
16 # 3D graph, scatter3d is from the --car package. It will open in the Viewer window
17
18 scatter3d(prestige ~ income + education, id.n=3, data=Duncan)
19 rglwidget()
```

Environment History Connections Tutorial

Import Dataset 359 MB Global Environment

Environment is empty

2 click on Zoom

1 Run

Files Plots Packages Help Viewer Presentation

Zoom Export

20:12 (Top Level) R Script

Console Terminal Background Jobs

R 4.3.0 - C:/Work/Data/

```
> scatterplot(prestige ~ income, data=Prestige)
> scatter3d(prestige ~ income + education, id.n=3, data=Duncan)
> rglwidget()
```

[See next slide]

OTR

Data visualization (3D graphics)

Use the cursor to move the graph around

