



UNIVERSITÀ  
DEGLI STUDI  
DI MILANO

HUMBOLDT-UNIVERSITÄT  
ZU BERLIN



# Session 00

## Installing the required software

---

Anette Fasang and Emanuela Struffolino

**Sequence Analysis for Social Science**

PDHP Workshop

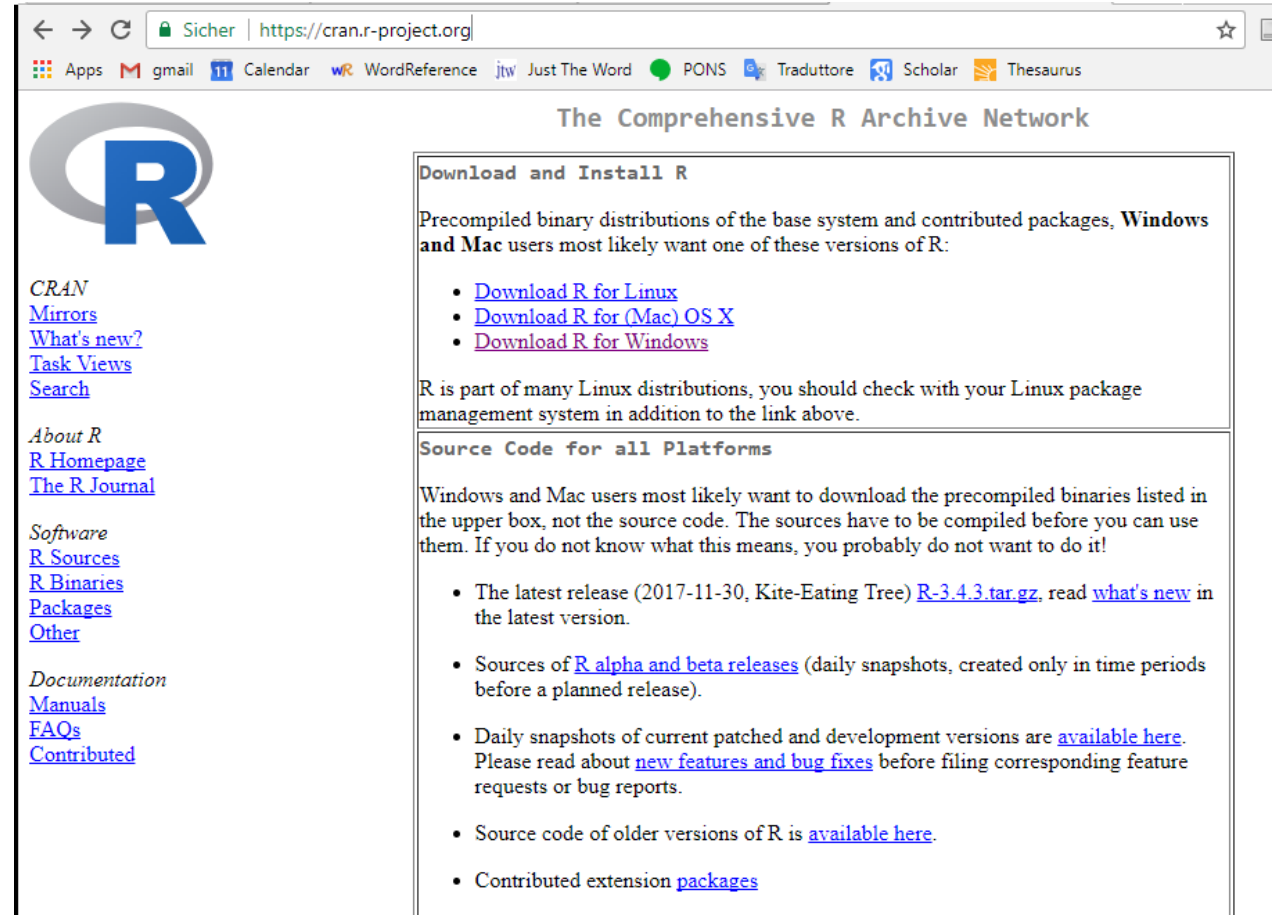
# Outline

- Installing R & Rstudio
- Installing R-packages
- Installing Graphviz
- Installing Stata-ados

# Installing R & RStudio

# Install R

- R is a free software environment for statistical computing and graphics
- Download:  
<https://cran.r-project.org/>
- Please download the latest version available on CRAN



The screenshot shows a web browser window displaying the CRAN website. The address bar shows the URL <https://cran.r-project.org/>. The page title is "The Comprehensive R Archive Network". The main content area is titled "Download and Install R" and contains the following text:

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux](#)
- [Download R for \(Mac\) OS X](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

**Source Code for all Platforms**

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

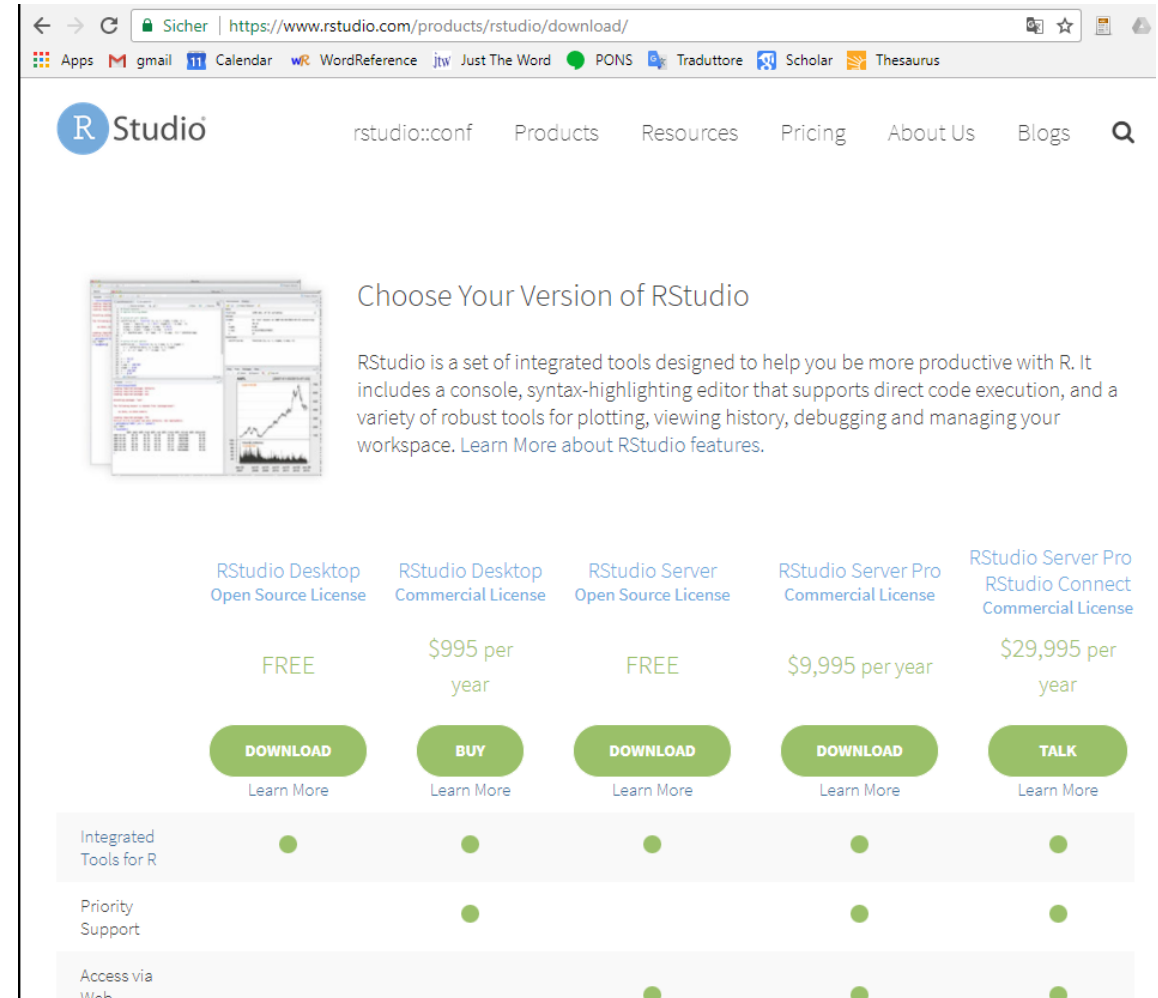
- The latest release (2017-11-30, Kite-Eating Tree) [R-3.4.3.tar.gz](#), read [what's new](#) in the latest version.
- Sources of [R alpha and beta releases](#) (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are [available here](#). Please read about [new features and bug fixes](#) before filing corresponding feature requests or bug reports.
- Source code of older versions of R is [available here](#).
- Contributed extension [packages](#)

The left sidebar of the website contains the following links:

- CRAN
- [Mirrors](#)
- [What's new?](#)
- [Task Views](#)
- [Search](#)
- About R
- [R Homepage](#)
- [The R Journal](#)
- Software
- [R Sources](#)
- [R Binaries](#)
- [Packages](#)
- [Other](#)
- Documentation
- [Manuals](#)
- [FAQs](#)
- [Contributed](#)

# Install RStudio

- RStudio is a free(ish) interface for R
- Download the FREE VERSION: <https://www.rstudio.com/>



The screenshot shows the RStudio website's download page. The browser address bar displays the URL <https://www.rstudio.com/products/rstudio/download/>. The page features a navigation menu with links for 'rstudio::conf', 'Products', 'Resources', 'Pricing', 'About Us', and 'Blogs'. A search icon is also present. Below the navigation, there is a section titled 'Choose Your Version of RStudio' with a small image of the RStudio interface. The text describes RStudio as a set of integrated tools for R, including a console, editor, and plotting tools. Below this, there are five columns representing different versions and licenses:

Version	License	Price	Action	Learn More
RStudio Desktop	Open Source License	FREE	DOWNLOAD	Learn More
RStudio Desktop	Commercial License	\$995 per year	BUY	Learn More
RStudio Server	Open Source License	FREE	DOWNLOAD	Learn More
RStudio Server Pro	Commercial License	\$9,995 per year	DOWNLOAD	Learn More
RStudio Server Pro RStudio Connect	Commercial License	\$29,995 per year	TALK	Learn More

Below the pricing table, there are three rows of features with green dots indicating availability:

Feature	Open Source	Commercial	Server Pro	Server Pro Connect
Integrated Tools for R	●	●	●	●
Priority Support		●		●
Access via Web			●	●

# Rstudio interface

- After installing RStudio, open it and have a look

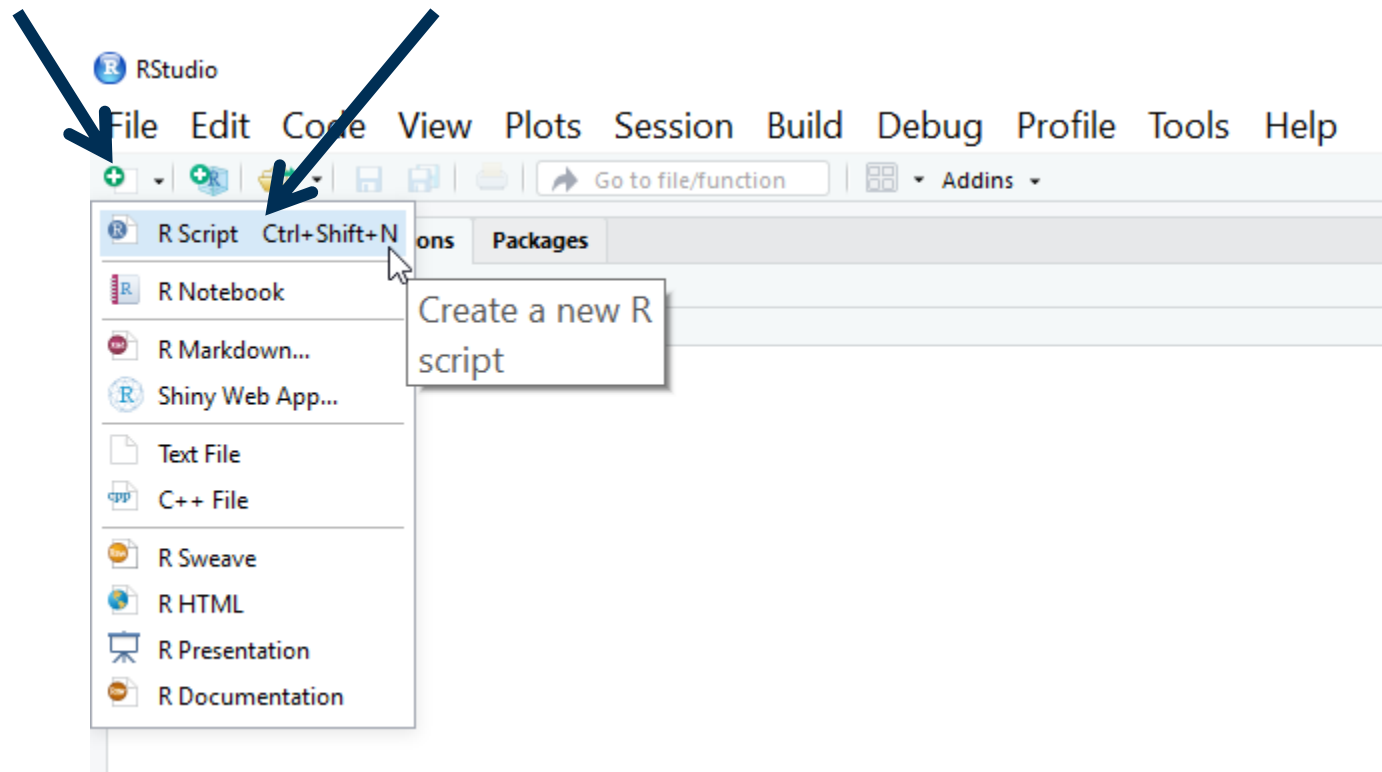
The screenshot displays the RStudio interface with three main panels highlighted by callouts:

- Console: output window**: The left panel shows the R version (3.4.2) and introductory text.
- Overview of objects in the environment (now empty)**: The top right panel shows the Environment pane, which is currently empty.
- Plots, help, packages...**: The bottom right panel shows the Packages pane, which lists installed and available packages.

Name	Description	Version
backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.0
base64enc	Tools for base64 encoding	0.1-3
BH	Boost C++ Header Files	1.65.0-1
bindr	Parametrized Active Bindings	0.1
bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2
bitops	Bitwise Operations	1.0-6
caTools	Tools: moving window statistics, GIF, Base64, ROC AUC, etc.	1.17.1

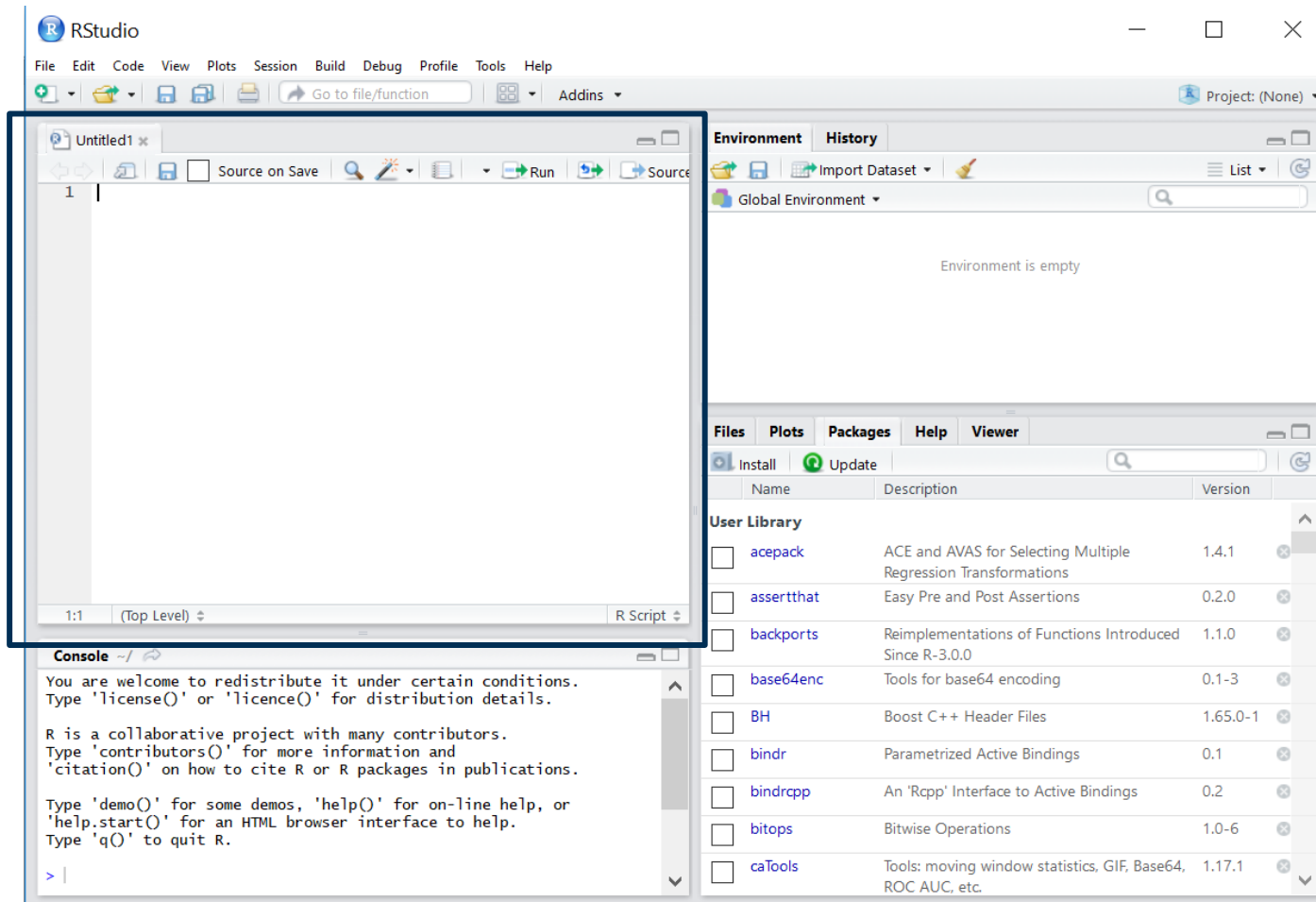
# Where do I write the commands

Click here and then here to open a new „R-script“ (command file)



# Where do I write the commands

... and you get this





# Installing R-packages

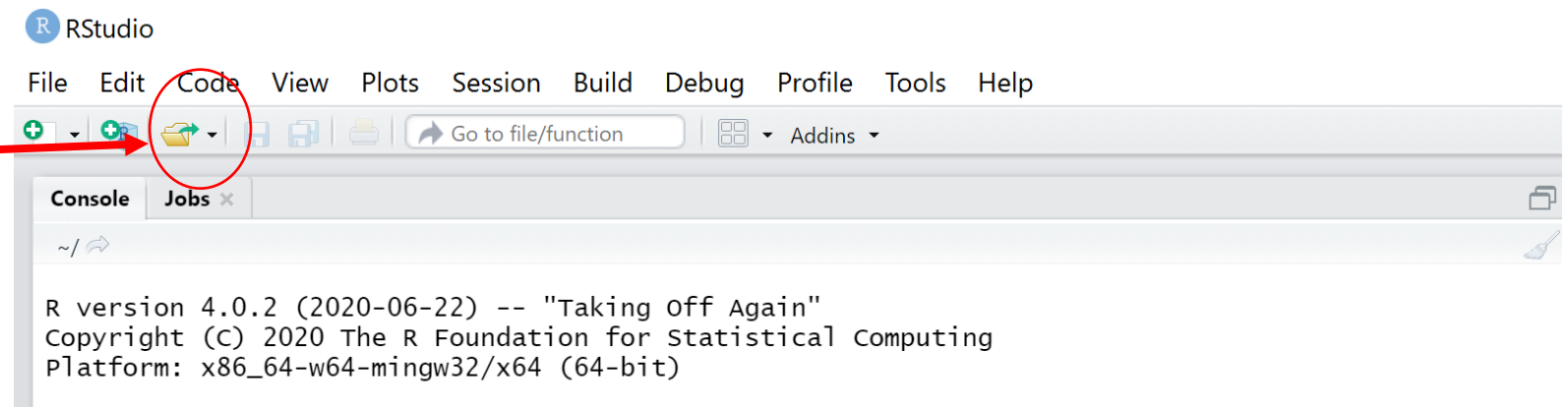
# Now install packages required for this course

- The R-packages are downloaded directly from the R-repository, and you need to an internet connection for that.
- Follow the steps on the next slides to open and run the shared R-Script **load\_and\_install\_packages.R**

# Installing R packages

Open R studio

Search for the file  
**load\_and\_install\_packages.R**  
here



You will get something like this:  
 colors in the window might  
 differ, no worries

```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
load_and_install_packages.R
#####
# Load and download (if necessary) packages in R-studio
#####
# this script illustrates an alternative way of installing and
# loading packages in R
# use (and install if necessary) pacman package
if (!require("pacman")) install.packages("pacman")
library(pacman)
# load and install (if necessary) required packages for this course
pacman::p_load(TraMineR, TraMineRextras, cluster, rio, ggplot2, plotrix,
              haven, Hmisc, ganimate, RColorBrewer, SDMTTools, colorspace,
              knitr, kableExtra, reshape2, summarytools, vegan, MCMCpack,
              corrplot, ade4, cssTools, weightedCluster, factoextra,
              dplyr, tidyverse, effects, margins, psych, devtools,
              broom, nnet, descr)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

1:1 (Untitled) R Script
Console Jobs
~/
R version 4.0.2 (2020-06-22) -- "Taking Off Again"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

Environment History Console
Global Environment
Files Plots Packages
R: Find statistics (including
statsBy {psych}
Find statistics (
between group:
Description
When examining data at tw
it is useful to find basic desc
correlations) as well as betw
between group correlations
correlations at the individua
groups.
Usage
statsBv(data, group.

```

# Installing R packages

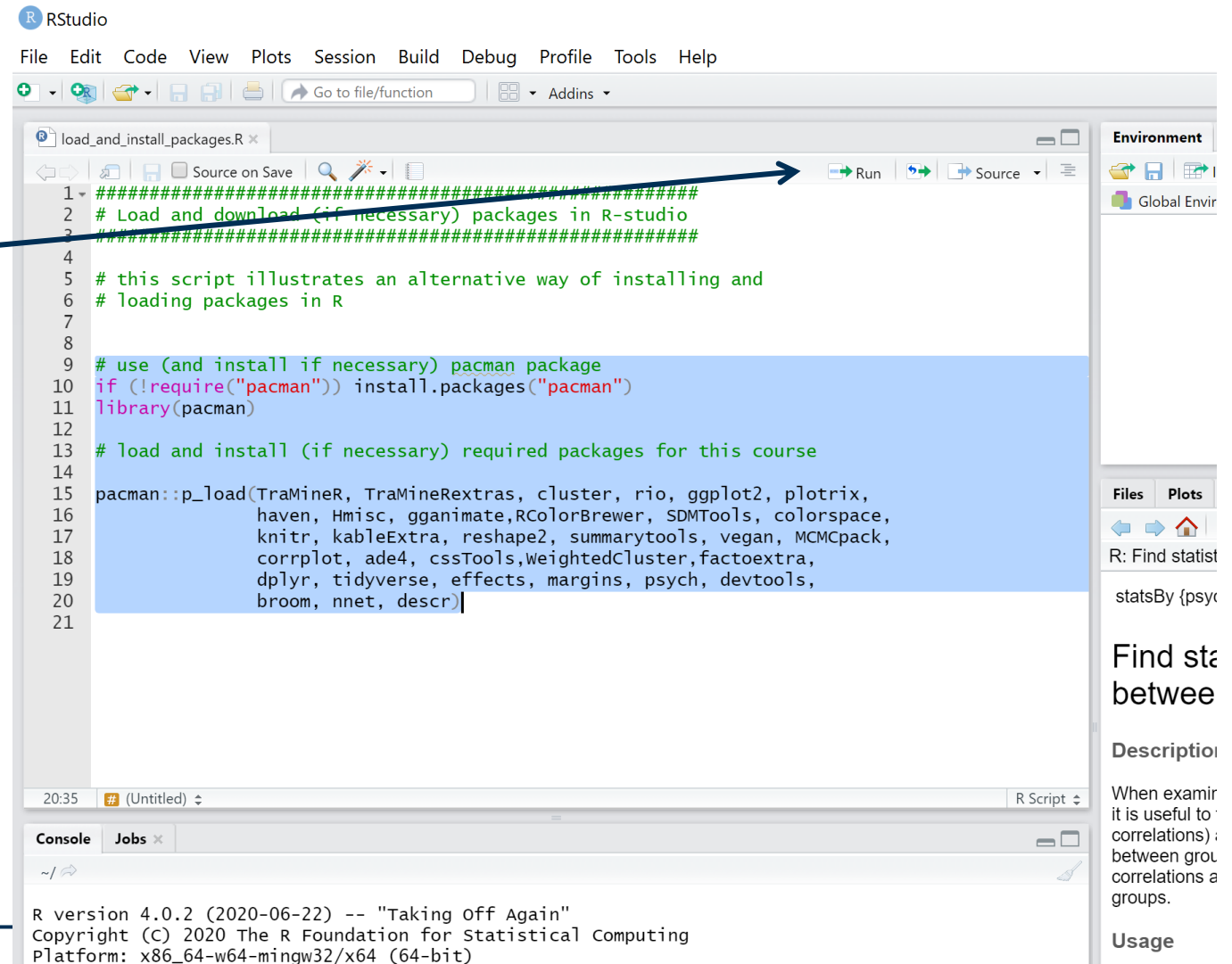
Now select the code as displayed (in light blue), and click the **Run**

or

use the shortcut

**Ctrl+Enter**

(shortcut might be different for you, especially if you are a Mac-OS user – but you can set them up, no worries)



The screenshot shows the RStudio interface with a script editor window titled 'load\_and\_install\_packages.R'. The script contains the following code:

```
1 #####  
2 # Load and download (if necessary) packages in R-studio  
3 #####  
4  
5 # this script illustrates an alternative way of installing and  
6 # loading packages in R  
7  
8  
9 # use (and install if necessary) pacman package  
10 if (!require("pacman")) install.packages("pacman")  
11 library(pacman)  
12  
13 # load and install (if necessary) required packages for this course  
14  
15 pacman::p_load(TraMineR, TraMineRextras, cluster, rio, ggplot2, plotrix,  
16 haven, Hmisc, gganimate, RColorBrewer, SDMTTools, colorspace,  
17 knitr, kableExtra, reshape2, summarytools, vegan, MCMCpack,  
18 corrrplot, ade4, cssTools, weightedCluster, factoextra,  
19 dplyr, tidyverse, effects, margins, psych, devtools,  
20 broom, nnet, descr)]  
21
```

The code from line 10 to line 21 is highlighted in light blue. A blue arrow points from the highlighted code to the 'Run' button in the top right corner of the script editor. The 'Run' button is also highlighted in light blue. The console at the bottom shows the R version and platform information:

```
R version 4.0.2 (2020-06-22) -- "Taking off Again"  
Copyright (C) 2020 The R Foundation for Statistical Computing  
Platform: x86_64-w64-mingw32/x64 (64-bit)
```

# Installing R packages

- You will see things going on in the console window, this means that R is downloading the packages from the repository and automatically installing them on your computer

- When R has finished, you should get something like this:

The screenshot shows the RStudio interface. The main editor window contains R code for installing several packages. The console window at the bottom shows the execution of this code, including the installation of the 'pacman' package and the loading of various packages like 'TraMineR', 'TraMineRextras', 'cluster', 'rio', 'ggplot2', and 'RColorBrewer'.

```

1 #####
2 # Download packages in R-studio
3 #####
4
5 # basic packages for Sequence Analysis
6
7 install.packages("TraMineR")
8 install.packages("TraMineRextras")
9 install.packages("WeightedCluster")
10 install.packages("cluster")
11
12 # packages for import/export data
13
14 install.packages("rio")
15
16 # packages for plotting
17
18 install.packages("plotrix")
19 install.packages("ggplot2")
20 install.packages("RColorBrewer")

```

```

> #####
> # Load and download (if necessary) packages in R-studio
> #####
>
> # this script illustrates an alternative way of installing and
> # loading packages in R
>
> # use (and install if necessary) pacman package
> if (!require("pacman")) install.packages("pacman")
> library(pacman)
>
> # load and install (if necessary) required packages for this course
>
> pacman::p_load(TraMineR, TraMineRextras, cluster, rio, ggplot2, plotrix,
+               haven, hmsc, gganimate, RColorBrewer, colorspace, MCMCpack,
+               knitr, kableExtra, reshape2, summarytools, vegan,
+               corrplot, ade4, cssTools, WeightedCluster, factoextra,
+               dplyr, tidyverse, effects, margins, psych, devtools,
+               broom, nnet, descr)
>

```

## Summary of the main packages

### # basic packages for Sequence Analysis

```
TraMineR  
TraMineRextras  
WeightedCluster  
cluster
```

### # packages for import/export data

```
rio  
haven
```

### # packages for plotting

```
plotrix  
ggplot2  
RColorBrewer  
colorspace  
ade4
```

# For now you're done with R!

- You can close the R window, the packages are saved on your computer!
- You don't need to save the code you typed in the Command window, we'll go back to that together during the GESIS workshop

BUT...

- ...there's one last R-related step to set up your computer:  
**installing Graphviz** (see next slide)



# Installing Graphviz

# How to install Graphviz

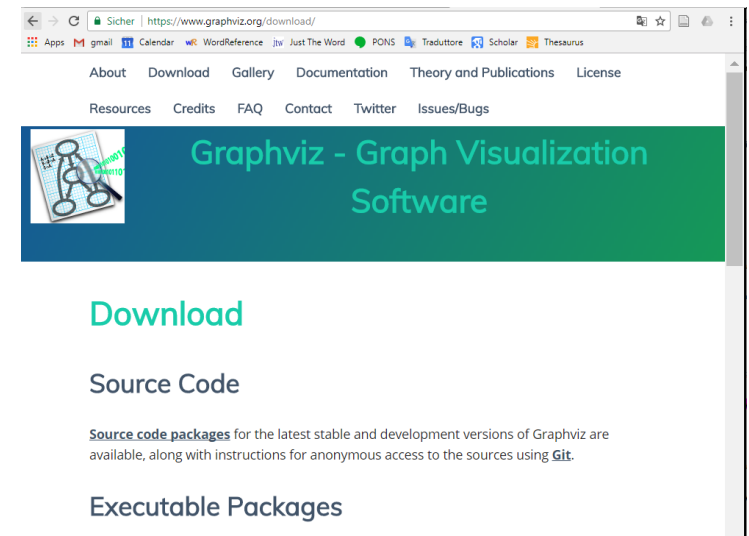
- Graphviz is a free Graph Visualization Tools that interacts with R for generating specific types of plots
- Download: <https://www.graphviz.org/download/>
- Executable Packages
  - Windows: Prefer the “Windows install packages” version, make sure you avoid the “development” version
  - Mac: MacPorts\*

← → ↻ [www2.graphviz.org/Packages/stable/windows/10/cmake/Release/x64/](https://www2.graphviz.org/Packages/stable/windows/10/cmake/Release/x64/)

Gmail 14 Google Calendar -... Survey Google Traduttore http://www.wordref... Engl Engl HU-Kalender

## Index of /Packages/stable/windows/10/cmake/Release/x64

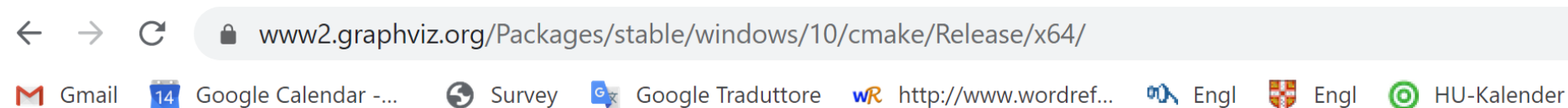
- [Parent Directory](#)
- [graphviz-install-2.44.1-win64.exe](#)



The screenshot shows the Graphviz website's download page. The browser address bar displays 'https://www.graphviz.org/download/'. The page features a navigation menu with links for 'About', 'Download', 'Gallery', 'Documentation', 'Theory and Publications', 'License', 'Resources', 'Credits', 'FAQ', 'Contact', 'Twitter', and 'Issues/Bugs'. A prominent green banner contains the text 'Graphviz - Graph Visualization Software' next to a small icon of a graph. Below the banner, there are sections for 'Download', 'Source Code', and 'Executable Packages'. The 'Source Code' section includes a note: 'Source code packages for the latest stable and development versions of Graphviz are available, along with instructions for anonymous access to the sources using Git.'

# How to install Graphviz

- You should follow these steps to get to the .exe file:



## Index of /Packages/stable/windows/10/cmake/Release/x64

- [Parent Directory](#)
- [graphviz-install-2.44.1-win64.exe](#)

Attention! Some recent updates of Graphviz/R generate problems so that although you manage to install Graphviz, R will not “find it” on your machine. This package is necessary only for regression trees in TraMineR, which probably we will not cover in this workshop. We wanted you to have this info, so that if you decide to use it in the future you are aware of this possible issue: unfortunately, so far the developers of the different packaged did not solve this problem structurally.

# Installing Stata-ados

We will not use Stata in this workshop, but - if you want to explore the sequence analysis option there – here are some instructions to get set!

# Install required Stata-ados

- We assume that you have Stata installed on your machines
- And that you are familiar with using do files
- For this course we need a couple of additional user written ados which can be installed by running the shared do-file (**InstallAdos.do**) or by copying and running the following commands

```
//SADI
net install st0486, from(http://www.stata-journal.com/software/sj17-3) replace

//lean scheme
net install gr0002_3, from(http://www.stata-journal.com/software/sj4-3) replace

//spost (requires version 13)
net install spost13_ado, from(https://jslsoc.sitehost.iu.edu/stata) replace

//graph combine
net install grc1leg, from(http://www.stata.com/users/vwiggins) replace

//Ados stored at ssc archive
local ados "distinct fre tsspell estout outreg2 sq moremata sadi blindschemes"
local ados "`ados' palettes"

foreach x of local ados {
    ssc install `x', replace
}
```

# Now you're done...

See you soon online!