

Unidad 1.3

R y Rstudio

- R
- Rstudio

Origen de R



- Desarrollado en 1992 por **Robert Gentleman** y **Ross Ihaka** en la Universidad de Auckland en Nueva Zelanda.
- Basado en el lenguaje de programación S desarrollado por Bell Company
- Distribuido como free software (GNU) en 1995



Robert Gentleman



Ross Ihaka

October 20, 2017

An Updated History of R

Here's a [refresher](#) on the history of the R project:

- 1992: R development begins as a research project in Auckland, NZ by Robert Gentleman and Ross Ihaka
- 1993: First binary versions of R published at Statlib [see update, below]
- 1995: R first distributed as open-source software, under GPL2 license
- 1997: R core group [formed](#)
- 1997: CRAN [founded](#) (by Kurt Jornik and Fritz Leisch)
- 1999: The R website, r-project.org, [founded](#)
- 1999: First in-person meeting of R Core team, at [inaugural Directions in Statistical Computing conference](#), Vienna
- 2000: R 1.0.0 [released](#) (February 29)
- 2000: John Chambers, [recipient](#) of the 1998 ACM Software Systems Award for the S language, [joins](#) R Core
- 2001: R News [founded](#) (later to become the R Journal)
- 2003: R Foundation [founded](#)
- 2004: First [UseR! conference](#) (in Vienna)
- 2004: R 2.0.0 [released](#)
- 2009: First edition of the R Journal
- 2013: R 3.0.0 [released](#)
- 2015: R Consortium [founded](#), with R Foundation participation
- 2016: New R logo [adopted](#)

<https://www.stat.auckland.ac.nz/~ihaka/downloads/Interface98.pdf>

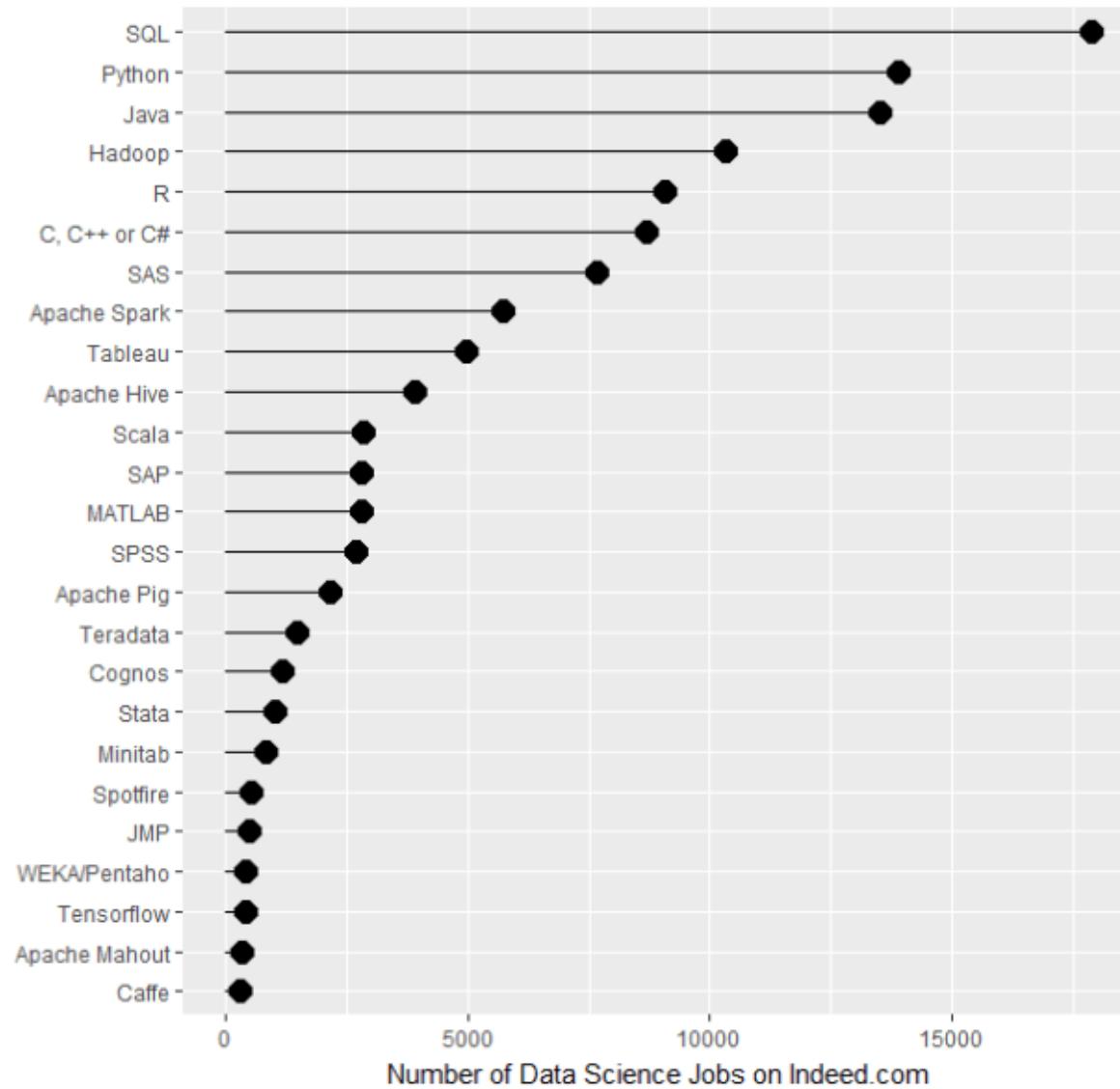
http://srv2.lemig.umontreal.ca/donnees/geo1512/IHAKA_JCGS_1996_R_concepts_implementati on.pdf

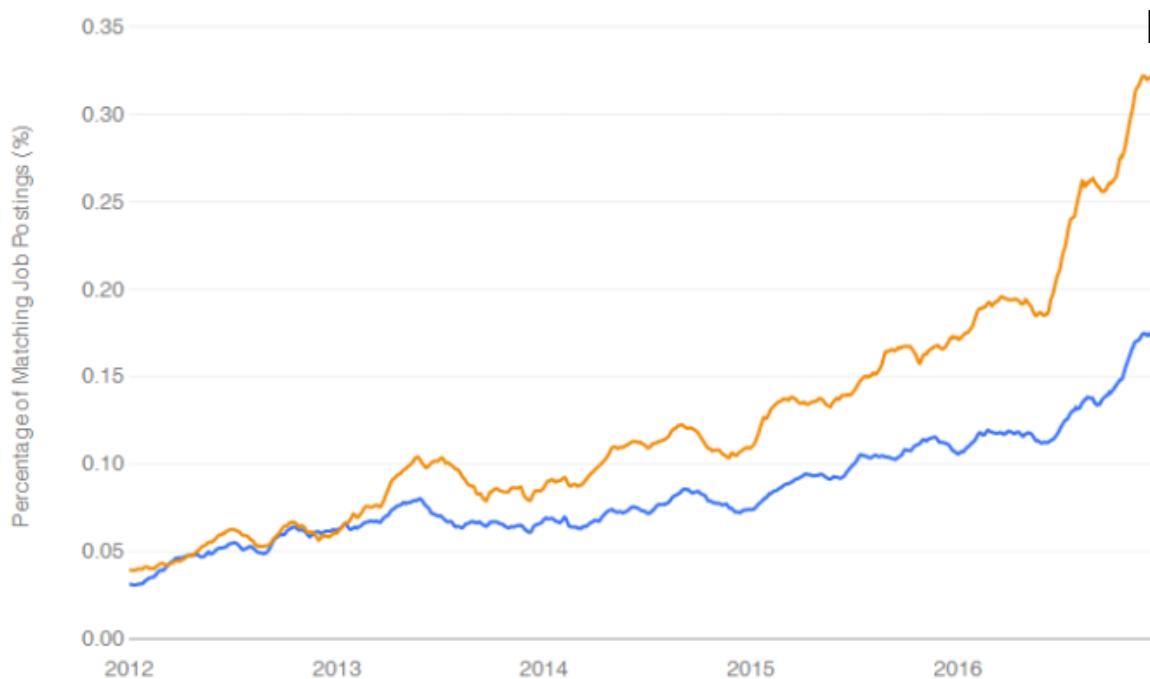
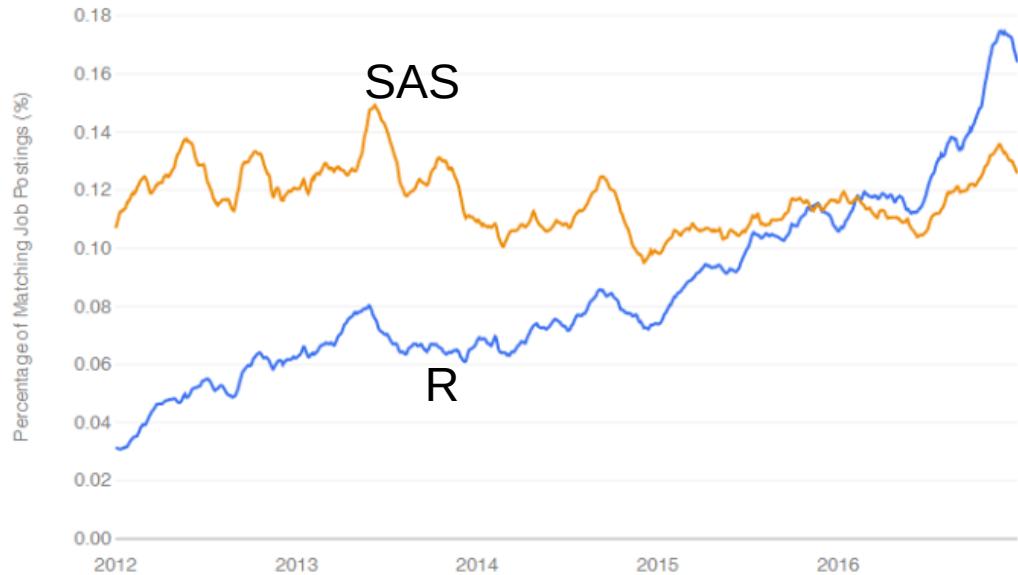


Among the goals of the R Foundation are the support of continued development of R, the exploration of new methodology, teaching and training of statistical computing and the organization of meetings and conferences with a statistical computing orientation.

R en el contexto de otros lenguajes

The More Popular Data Science Software (>=250 Jobs)





Pero tienen fortalezas diferentes!

Python lenguaje de programación de fácil lectura

R análisis estadístico

Ver link
<https://bit.ly/2Xh9ERi>

Descripción general de



- Lenguaje y ambiente de programación (interpreter)
- Diseñado para exploración y análisis de datos y puede crecer gracias a la implementación de paquetes especializados en diferentes tipos de análisis o procedimientos.
- Página oficial <http://www.r-project.org>
- Los paquetes creados por la comunidad están almacenados en CRAN - Comprehensive R Archive Network, y otros como BioConductor , omegahat , RForge.
- Nuevas versiones completas son generadas cada año
- Puede ser usado con o sin un Integrated Development Environment (IDE) como Rstudio
- No necesariamente bueno para cálculos intensivos (number crunching)

[\[Home\]](#)

Download

[CRAN](#)

R Project

[About R](#)[Logo](#)[Contributors](#)[What's New?](#)[Reporting Bugs](#)[Conferences](#)[Search](#)[Get Involved: Mailing Lists](#)[Developer Pages](#)[R Blog](#)

R Foundation

[Foundation](#)[Board](#)[Members](#)[Donors](#)[Donate](#)

Help With R

[Getting Help](#)

Documentation

[Manuals](#)[FAQs](#)[The R Journal](#)[Books](#)[Certification](#)[Other](#)

Links

[Bioconductor](#)[Related Projects](#)[GSoC](#)

The R Project for Statistical Computing

Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To [download R](#), please choose your preferred CRAN mirror.

If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

News

- [R version 3.6.0 \(Planting of a Tree\)](#) has been released on 2019-04-26.
- useR! 2020 will take place in St. Louis, Missouri, USA.
- [R version 3.5.3 \(Great Truth\)](#) has been released on 2019-03-11.
- The R Foundation Conference Committee has released a call for proposals to host useR! 2020 in North America.
- You can now support the R Foundation with a renewable subscription as a [supporting member](#)
- The R Foundation has been awarded the Personality/Organization of the year 2018 award by the professional association of German market and social researchers.

News via Twitter

 **The R Foundation** @_R_Foundation

New R blog entry by Tomas Kalibera explaining recent problems with Fortran.[developer.r-project.org/Blog/public/20...](#)



May 15, 2019

 **The R Foundation** @_R_Foundation

Replies to @_R_Foundation
Problems with gfortran 9.0 may be mitigated by disabling tail-call optimization (option -fno-optimize-sibling-calls).



May 3, 2019

Mirrors

Secure https://cran.r-project.org/mirrors.html	
https://cran.hafro.is/	Marine Research Institute
http://cran.hafro.is/	Marine Research Institute
Indonesia	
https://repo.bppr.go.id/cran/	Agency for The Application and Assessment of Technology
Iran	
https://cran.um.ac.ir/	Ferdowsi University of Mashhad
http://cran.um.ac.ir/	Ferdowsi University of Mashhad
Ireland	
https://ftp.heanet.ie/mirrors/cran.r-project.org/	HEAnet,Dublin
http://ftp.heanet.ie/mirrors/cran.r-project.org/	HEAnet,Dublin
Italy	
http://cran.mirror.garr.it/mirrors/CRAN/	Garr Mirror, Milano
https://cran.stat.unipd.it/	University of Padua
http://cran.stat.unipd.it/	University of Padua
Japan	
https://cran.ism.ac.jp/	The Institute of Statistical Mathematics, Tokyo
http://cran.ism.ac.jp/	The Institute of Statistical Mathematics, Tokyo
https://ftp.yz.yamagata-u.ac.jp/pub/cran/	Yamagata University
Korea	
https://ftp.harukasan.org/CRAN/	Information and Database Systems Laboratory, Pukyong National University
https://cran.yu.ac.kr/	Yeungnam University
https://cran.seoul.go.kr/	Bigdata Campus, Seoul Metropolitan Government
http://healthstat.snu.ac.kr/CRAN/	Graduate School of Public Health, Seoul National University, Seoul
https://cran.biostat.org/	The Genome Institute of UNIST (Ulsan National Institute of Science and Technology)
http://cran.biostat.org/	The Genome Institute of UNIST (Ulsan National Institute of Science and Technology)
Malaysia	
https://wbc.upm.edu.my/cran/	Universiti Putra Malaysia
http://wbc.upm.edu.my/cran/	Universiti Putra Malaysia
Mexico	
https://cran.itam.mx/	Instituto Tecnologico Autonomo de Mexico
http://cran.itam.mx/	Instituto Tecnologico Autonomo de Mexico
http://www.est.colpos.mx/R-mirror/	Colegio de Postgraduados, Texcoco
New Zealand	
https://cran.stat.auckland.ac.nz/	University of Auckland
http://cran.stat.auckland.ac.nz/	University of Auckland
Norway	
https://cran.uib.no/	University of Bergen
http://cran.uib.no/	University of Bergen
Philippines	
https://cran.stat.upd.edu.ph/	University of the Philippines and PREGINET
http://cran.stat.upd.edu.ph/	University of the Philippines and PREGINET
Portugal	
http://cran.radicaldevelop.com/	RadicalDevelop, Lda
http://cran.dcc.fc.up.pt/	University of Porto

Everything in R is an object



objeto



Clase A

Generic function

Método

- Atributos
(ej. 4 patas y respaldo)

Silla

sientate

- Sentarte
- Subirte
- Atrancar puerta



iris

A	B	C
1	a	8
3	b	6
2	a	7
1	b	4

- Atributos
(ej. 4 patas y respaldo)

Silla

sientate

- Sentarte
- Subirte
- Atrancar puerta

iris

A	B	C
1	a	8
3	b	6
2	a	7
1	b	4

- Atributos
(ej. longitud, tipo de valores)

numeric

mean()

factor

levels()

dim.data.frame()

mean.Date()

mean.default()

mean.difftime()

mean.POSIXct()

mean.POSIXlt()

mean.quosure()

levels.default()

Toda clase tiene un método que es asignado por una función genérica (el Método es despachado por una función genérica)

methods(mean)

Class: describe objetos

Objetos: son casos de una clase

Method: función que solo opera en ciertas clases de objetos

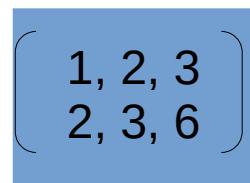
Generic function: asigna (despacha) a un objeto una función particular a su clase, i.e. un method

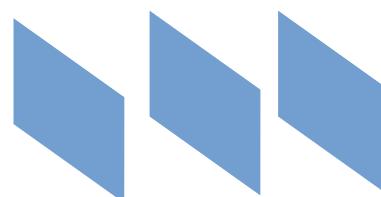


1 Numeric

c(1, 2, 3) Numeric (vector)

list(1, 2, 3, "holá") list

 matrix

 array

trat	Alt.	w
A	123	5
A	223	3
B	34	5

.

.

.

fit<-lm(A~B) lm

.

.

.

Mi propia clase

En R prueba

```
class (1)
```

```
class(c(1, 2, 3))
```

```
class(list(1,2,3, "holá"))
```

```
class (matrix(as.matrix(1:10),  
Nrow = 2))
```

```
class (array(1:24, dim=c(3,4,2)))
```

```
class (CO2)
```

```
class (mean)
```

```
# Creando un modelo lineal (lm) para mostrar dif tipos de
# objetos

tratamiento <- as.factor(c("w","w", "p", "p"))
altura <- as.numeric(c(4.2,5.4,2.1,3.2))
mi_base <- data.frame(tratamiento, altura)

lm_result <- lm(altura ~ tratamiento) # corriendo modelo lineal
anova_res <- anova(lm_result) # salvando anova

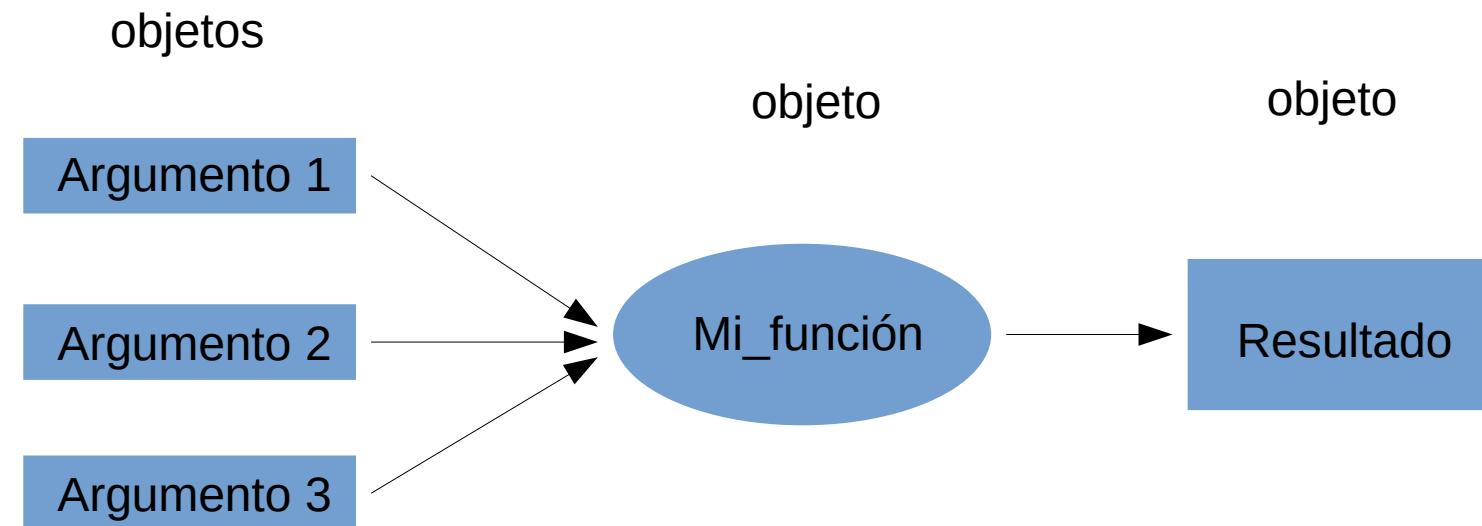
str(lm_result) # viendo estructura del output
class (lm_result)# preguntando por la clase del objeto

str(anova_res)
class(anova_res)
```

Si corres la rutina, verás que `lm_result` y `anova_res` son objetos de clases diferentes y que son tratadas de manera diferente.

R programación de objetos

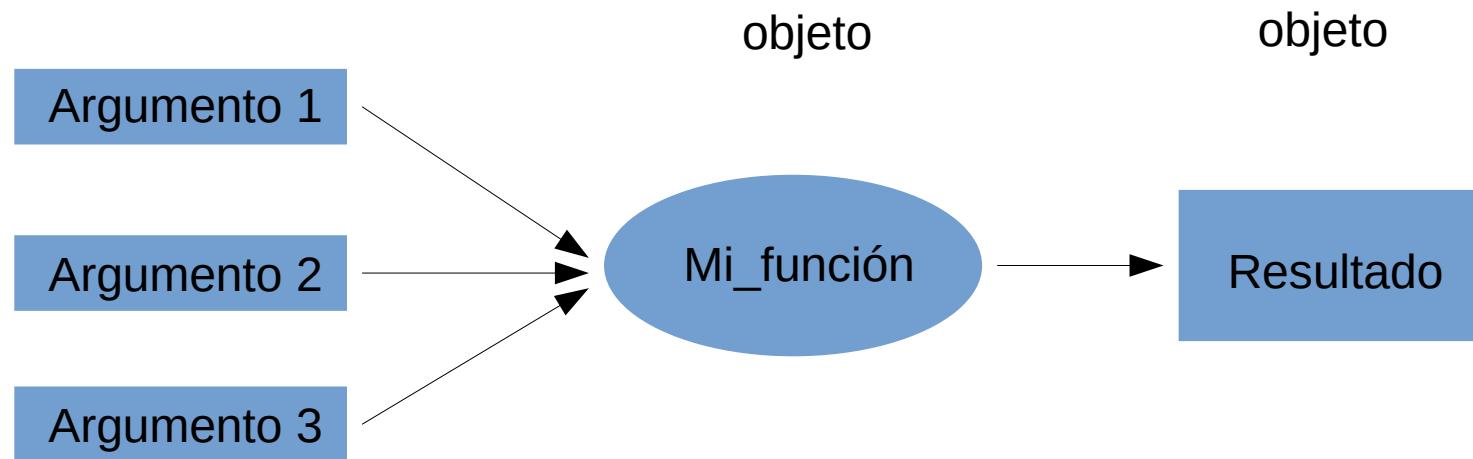
```
lm_result <- lm(altura ~ tratamiento, mi_base)
```



R programación de objetos

```
lm_result <- lm(altura ~ tratamiento, mi_base)
```

objetos



```
mi_funcion <- function (x, y, z) {  
  w <- x + y + z  
  print (w)  
}
```

```
mi_resultado <- mi_funcion(argumento1, argumento2, argumento3)
```

Funciones como objetos

```
mi_funcion <- function (x, y, z) {  
    # acepta 3 numeric obj  
    # regresa las suma  
    w <- x + y + z  
    print (w)  
}  
  
mi_resultado <- mi_funcion(argumento1, argumento2, argumento3)  
  
mi_resultado <- mi_funcion(5, 5, 5)  
> 15  
class (mi_funcion) #
```

Dos tipos de Clases y Métodos

- S3
 - Heredado del language S
 - Informal y obscuro
 - Llamado a veces old-style classes/methods
 - No permite la incorporación de nuevas clases
 - Usa \$ para referir una variable (A\$mi_objeto)
- S4
 - Más formal y riguroso
 - Llamado a veces new-style classes/methods
 - Permite incorporación de nuevas clases
 - Usa @ para referir una variable (A@mi_objeto)

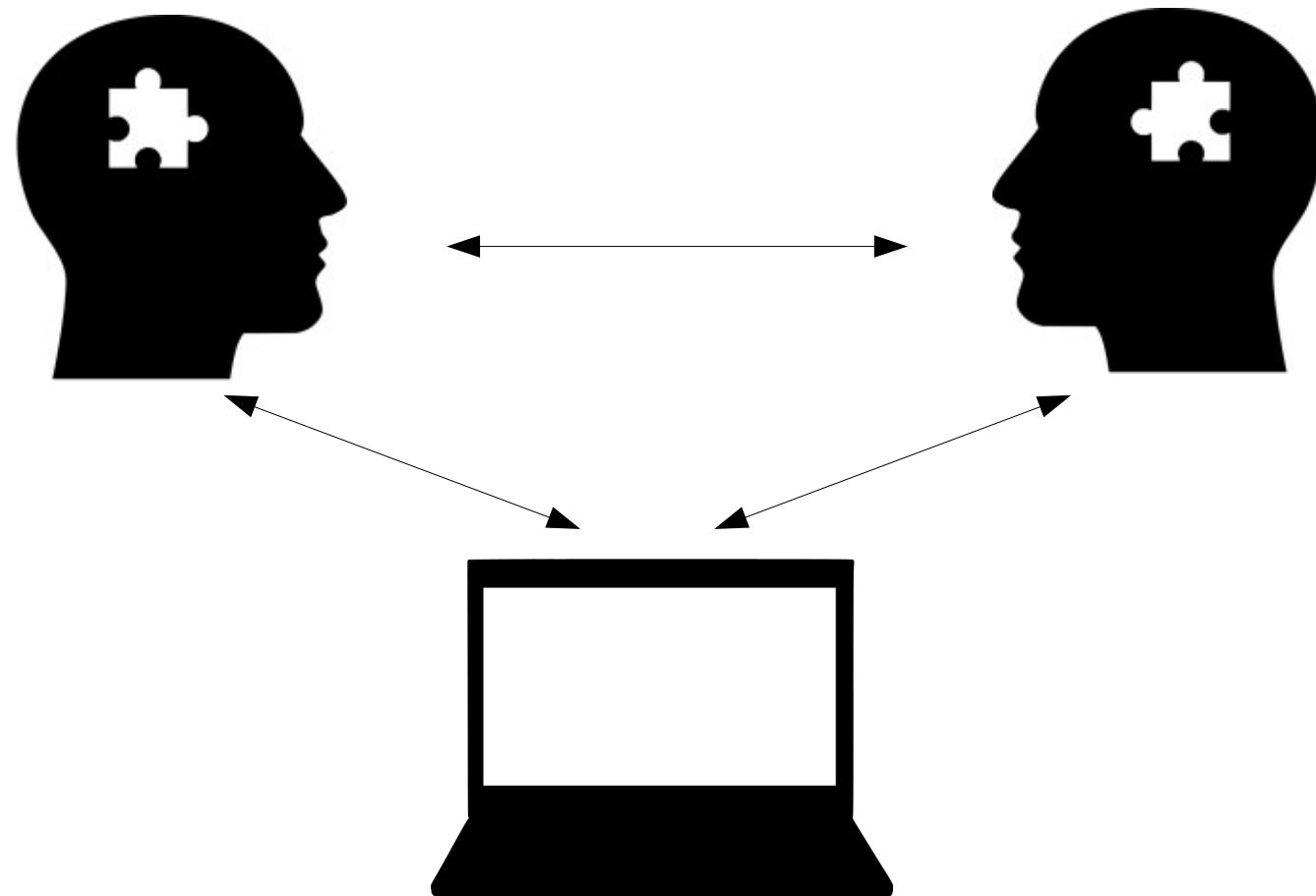
Guías de estilo de R

R es un lenguaje y requiere estilo (formato) para ser entendido



Guías de estilo de R

R es un lenguaje y requiere estilo (formato) para ser entendido



Google's R style Guide

<https://google.github.io/styleguide/Rguide.xml>

R style Guide by Hadley Wickham



ggplot2
tidyverse
Rstudio

<http://adv-r.had.co.nz/Style.html>

Nombres de files
Good
fit-models.R
utility-functions.R

Bad
foo.r
stuff.r

SE CONSISTENTE!

Nombres de objetos
Good
day_one
day_1

Bad
first_day_of_the_month
DayOne
dayone
djm1

The State of Naming Conventions in R

https://journal.r-project.org/archive/2012-2/RJournal_2012-2_Baaaath.pdf

Introducción a Rstudio

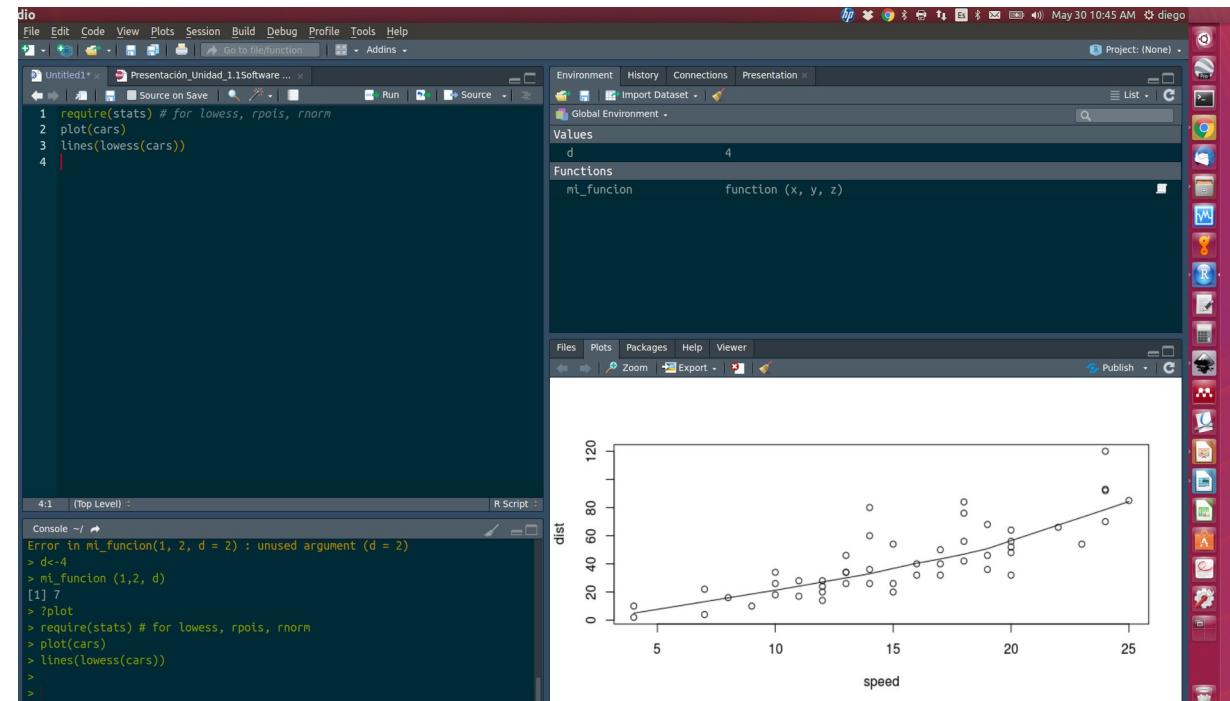
<https://support.rstudio.com/hc/en-us>



R en terminal en linux

```
o@cosimo:~  
diego@cosimo:~$ R  
  
R version 3.4.4 (2018-03-15) -- "Someone to Lean On"  
Copyright (C) 2018 The R Foundation for Statistical Computing  
Platform: x86_64-pc-linux-gnu (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.  
  
Natural language support but running in an English locale  
  
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.  
  
[Previously saved workspace restored]  
  
> print ("hola mundo")  
[1] "hola mundo"  
>
```

R usando un Integrated Development Environment (IDE)





RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Untitled1* x Presentación_Unidad_1.1Software ...

Source on Save Run Source

1 a<-c(1:10)
2 b<-list(1:10)
3 c <- matrix(c(1,2,3, 11,12,13), nrow = 2, ncol = 3, byrow = TRUE,
4 dimnames = list(c("row1", "row2"),
5 c("C.1", "C.2", "C.3")))
6
7 d<- data.frame("Name" = c("Sp1","Sp2"), "Sample" = 1:2, "Age" = c(21, 15))
8 e<-c("hola", "mundo")
9

Source

9:1 (Top Level) R Script

Console ~/

```
> a<-c(1:10)
> b<-list(1:10)
> c <- matrix(c(1,2,3, 11,12,13), nrow = 2, ncol = 3, byrow = TRUE,
+ dimnames = list(c("row1", "row2"),
+ c("C.1", "C.2", "C.3")))
>
> d<- data.frame("Name" = c("Sp1", "Sp2"), "Sample" = 1:2, "Age" = c(21, 15))
> e<-c("hola", "mundo")
>
```

Console

Environment

Data

- b List of 1 : int [1:10] 1 2 3 4 5 6 7 8 9 10
- c num [1:2, 1:3] 1 11 2 12 3 13
- d 2 obs. of 3 variables Name : Factor w/ 2 levels "Sp1","Sp2": 1 2 Sample: int 1 2 Age : num 21 15

Values

- a int [1:10] 1 2 3 4 5 6 7 8 9 10
- e chr [1:2] "hola" "mundo"

Files Plots Packages Help Viewer

R Data Frames Find in Topic

data.frame {base} R Documentation

Files, plots, etc.

Data Frames

Description

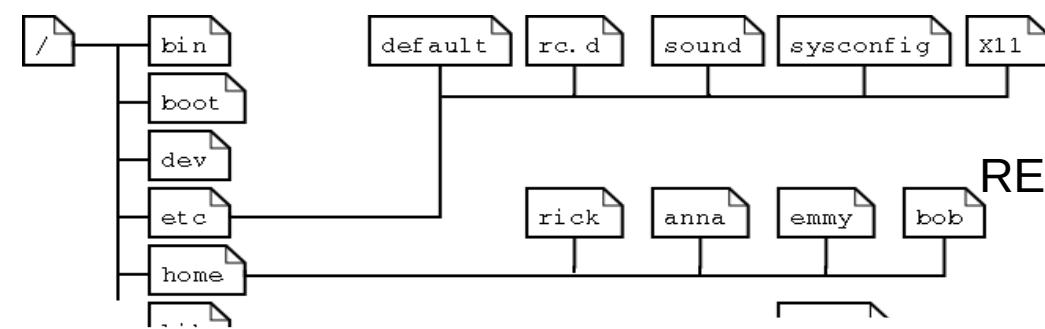
The function `data.frame()` creates data frames, tightly coupled collections of variables which share many of the properties of matrices and of lists, used as the fundamental data structure by most of R's modeling software.

Usage

```
data.frame(..., row.names = NULL, check.rows = FALSE,  
check.names = TRUE, fix.empty.names = TRUE,  
stringsAsFactors = default.stringsAsFactors())  
  
default.stringsAsFactors()
```

Arguments

Directarios de trabajo



Manten todos los files de los proyecto juntos

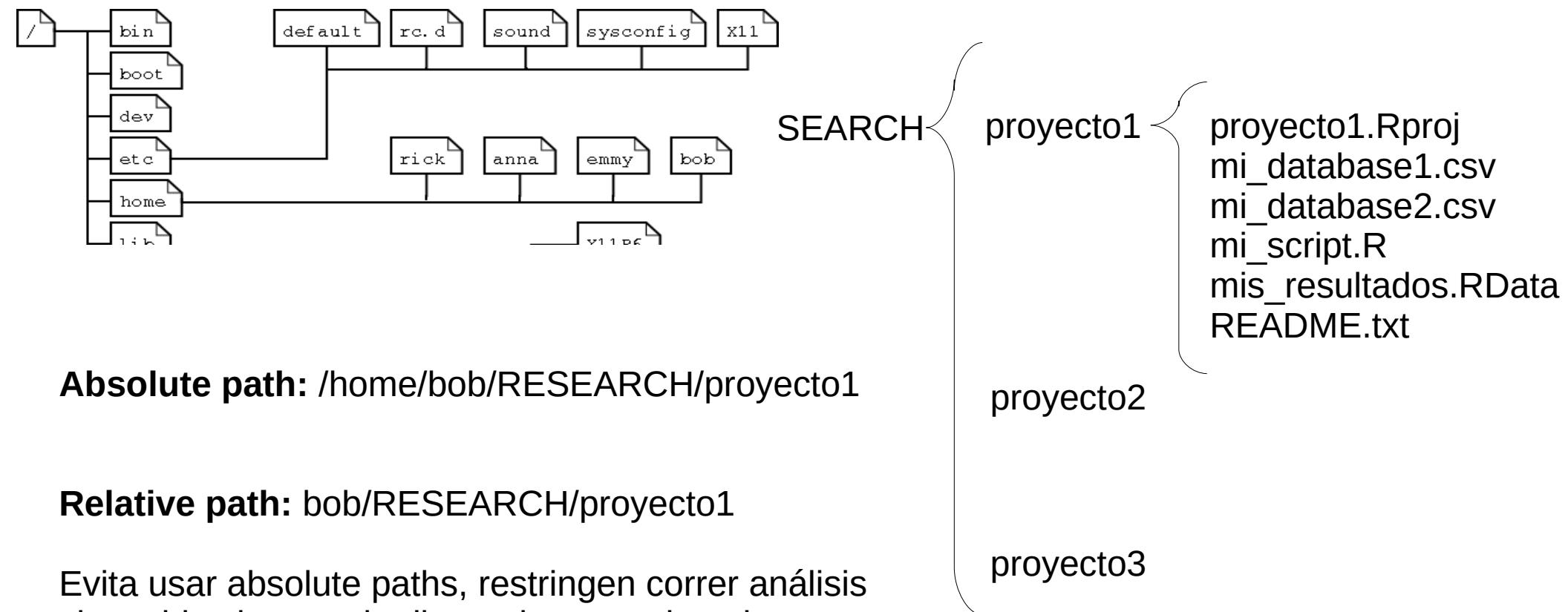
RESEARCH {

- proyecto1 {
- proyecto1.Rproj
- mi_database1.csv
- mi_database2.csv
- mi_script.R
- mis_resultados.RData
- README.txt

proyecto2

proyecto3

Directarios de trabajo



10

The screenshot shows a file browser window with a sidebar on the left containing 'Recientes' (Recent), 'Favoritos' (Favorites), 'Carpeta personal' (Personal Folder), and 'Escritorio' (Desktop). The main area displays two tabs: 'cursos' and 'Unidad_1_Introduccion_general'. The 'cursos' tab is active, showing a list of items with columns for Nombre, Tamaño, and Modificación. The first item is 'curso 2021 I' (4 elementos, 17:08) and the second is 'Introduccion_a_Manejo_y_Analisis_de_datos_en_R' (6 elementos, Ayer).

Nombre	Tamaño	Modificación
curso 2021 I	4 elementos	17:08
Introduccion_a_Manejo_y_Analisis_de_datos_en_R	6 elementos	Ayer

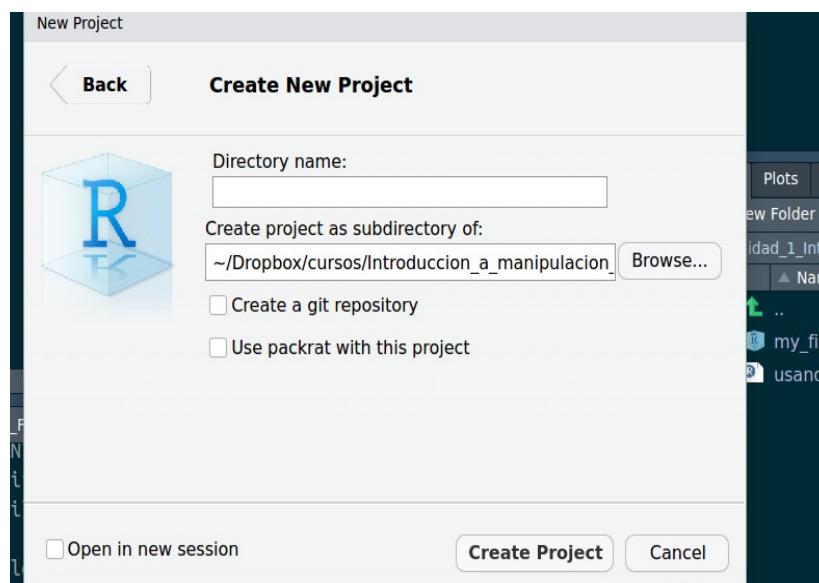
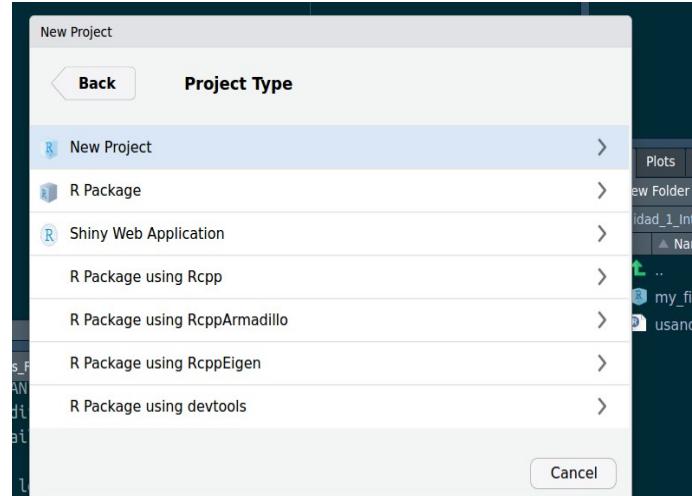
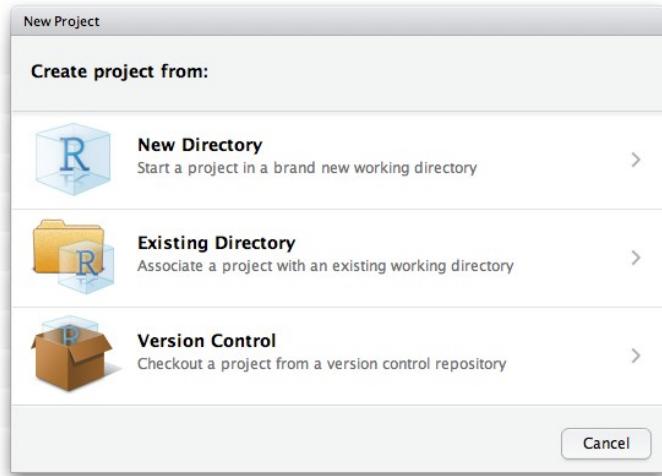
10

The screenshot shows a file browser window with a sidebar on the left containing 'Recientes' (Recent), 'Favoritos' (Favorites), 'Carpeta personal' (Personal Folder), 'Escritorio' (Desktop), and 'Documents'. The main area displays four tabs: 'Introdu...', 'Introdu...', 'Introdu...', and 'Unidad...'. The 'Unidad...' tab is active, showing a list of items with columns for Nombre, Tamaño, and Modificación. The items listed are: 'Unidad_1_Introduccion_general' (8 elementos, Ayer), 'Unidad_2_Manage_base_datos' (0 elementos, Ayer), 'Unidad_3_Vectores_listas_matrices_arrays' (0 elementos, Ayer), 'Unidad_4_Funciones_y_loops' (0 elementos, Ayer), 'Unidad_5_Graficos' (0 elementos, Ayer), and 'Unidad_6_Modelos' (0 elementos, Ayer).

Nombre	Tamaño	Modificación
Unidad_1_Introduccion_general	8 elementos	Ayer
Unidad_2_Manage_base_datos	0 elementos	Ayer
Unidad_3_Vectores_listas_matrices_arrays	0 elementos	Ayer
Unidad_4_Funciones_y_loops	0 elementos	Ayer
Unidad_5_Graficos	0 elementos	Ayer
Unidad_6_Modelos	0 elementos	Ayer

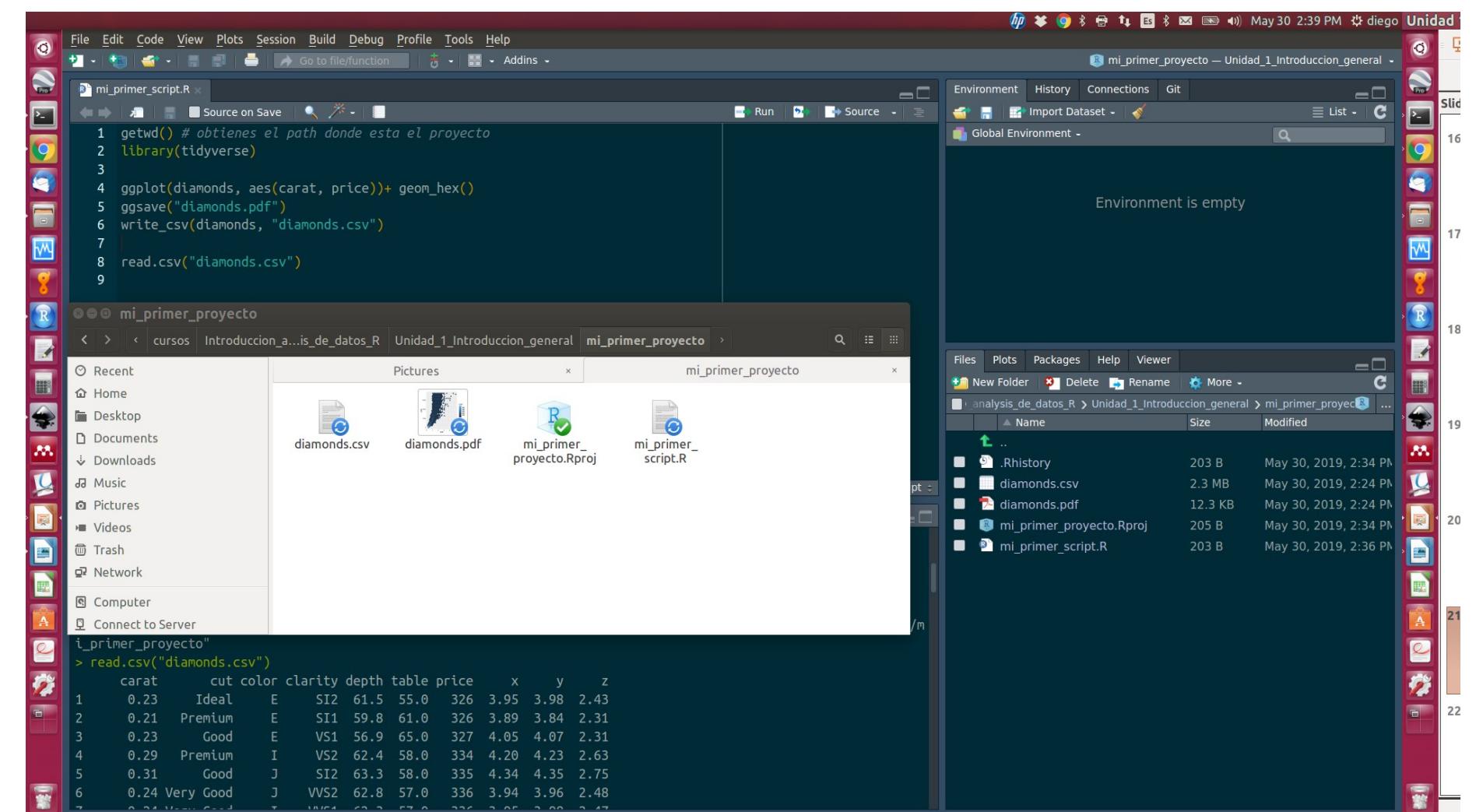
The screenshot shows a file browser window with a sidebar on the left containing 'Recientes' (Recent), 'Favoritos' (Favorites), 'Carpeta personal' (Personal Folder), 'Escritorio' (Desktop), and 'Documents'. The main area displays five tabs: 'Introdu...', 'Unidad_1_Introduccion_general', 'Unidad_1_Introduccion_general', 'Unidad_1_Introduccion_general', and 'Unidad_1_Introduccion_general'. The third tab is active, showing a list of items with columns for Nombre, Tamaño, and Modificación. The items listed are: 'database' (2 elementos, Ayer), 'literatura_soporte' (4 elementos, Ayer), 'presentaciones' (9 elementos, 20:24), 'src' (1 elemento, Ayer), '.Rproj.user' (2 elementos, Ayer), 'README' (452 bytes, Ayer), 'Unidad_1.Rproj' (205 bytes, Ayer), and '.Rhistory' (927 bytes, Ayer).

Nombre	Tamaño	Modificación
database	2 elementos	Ayer
literatura_soporte	4 elementos	Ayer
presentaciones	9 elementos	20:24
src	1 elemento	Ayer
.Rproj.user	2 elementos	Ayer
README	452 bytes	Ayer
Unidad_1.Rproj	205 bytes	Ayer
.Rhistory	927 bytes	Ayer



Cuando se **crea** un proyecto:

- 1) un file .Rproj es creado
- 2) un file .Rproj.user creado (oculto) que guarda files temporales
- 3) Se carga el proyecto nuevo en Rstudio.
- 4) Se recomienda agregar un README file



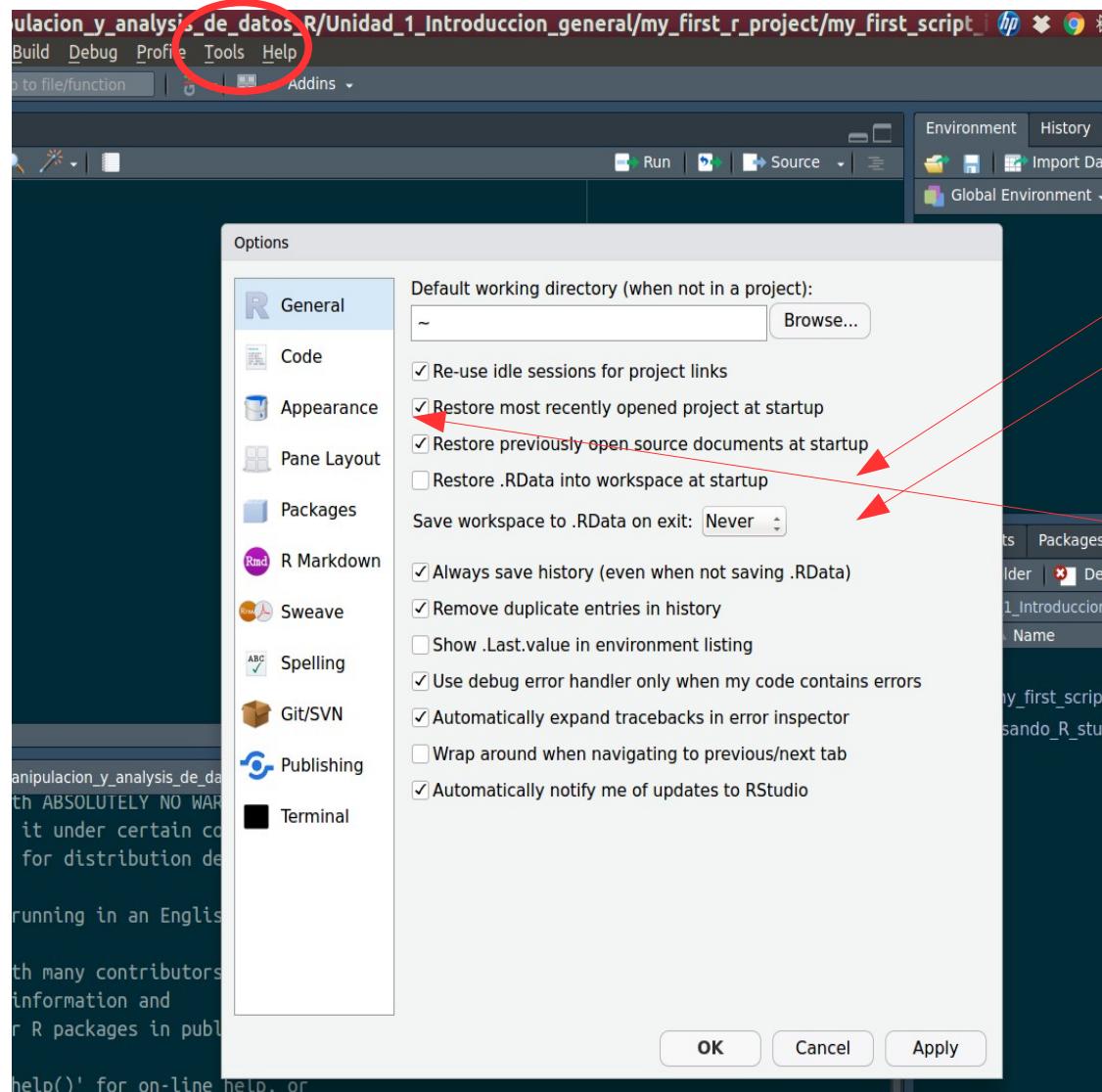
Cuando se abre el proyecto

- 1) Inicia una nueva sesión de R
- 2) Se carga el .Rprofile preferencias Rstudio
- 3) Se carga el .Rdata file en el directorio del proyecto
- 4) Se carga la historia
- 5) Se establece el directorio actual como el directorio de trabajo

Cuando se cierra el proyecto (quit, close, open another p.)

- 1) .Rdata y/o .Rhistory son escritos en el dir del proyecto (si así se establece en los settings)
- 2) La lista de documentos de open source son salvados
- 3) Otros settings de Rstudio son salvados
- 4) La sesión se da por terminada

Ajustes iniciales a Rstudio



En herramientas/tools

General options

Desactivar Restore .Rdata

Deactivar save workspace

En appearance puedes manipular
Color de pantalla, tamaño de letra
Etc.

Paquetes

Instalando el paquete Ime4

The screenshot shows the RStudio interface. On the left, the R console displays the standard R startup message. In the center, an 'Install Packages' dialog box is open, prompting the user to install 'ime4' from CRAN. To the right, the Global Environment pane shows an empty environment. A red circle highlights the 'Install' button in the Packages tab of the tools bar.

Install Packages

Install from: Repository (CRAN)

Packages (separate multiple with space or comma): **ime4**

Install to Library: /home/diego/R/x86_64-pc-linux-gnu-library/3.4 [Default]

Install dependencies

Install Cancel

Global Environment

Environment is empty

Packages

Name	Description	Version
abind	Combine Multidimensional Arrays	1.4-5
acepack	ACE and AVAS for Selecting Multiple Regression Transformations	1.4.1
ade4	Analysis of Ecological Data: Exploratory and Euclidean Methods in Environmental Sciences	1.7-13
animation	A Gallery of Animations in Statistics and Utilities to Create Animations	2.6
AnnotationDbi	Annotation Database Interface	1.40.0
ape	Analyses of Phylogenetics and Evolution	5.2
askpass	Safe Password Entry for R, Git, and SSH	1.1
assertthat	Easy Pre and Post Assertions	0.2.0
backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.2
base64enc	Tools for base64 encoding	0.1-3
BB	Solving and Optimizing Large-Scale Nonlinear Systems	2014.10-1
BDgraph	Bayesian Structure Learning in Graphical Models using Birth-Death MCMC	2.55
BH	Boost C++ Header Files	1.69.0-1
bindr	Parametrized Active Bindings	0.1.1

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. The toolbar contains icons for file operations like Open, Save, and Print, along with Go to file/function, Source on Save, and Addins. A code editor window titled 'Untitled1*' contains the following R code:

```
1 library (lme4)
2
```

The status bar at the bottom indicates the current workspace size as 2:1 and the path as (Top Level). The R console output below the code editor shows the standard R startup message, followed by the command and its execution:

```
Console ~/Dropbox/cursos/introduccion_a_manipulacion_y_analysis_de_datos_R/Unidad_1_Introduccion_general
Platform: x86_64-pc-linux-gnu (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.

Natural language support but running in an English locale

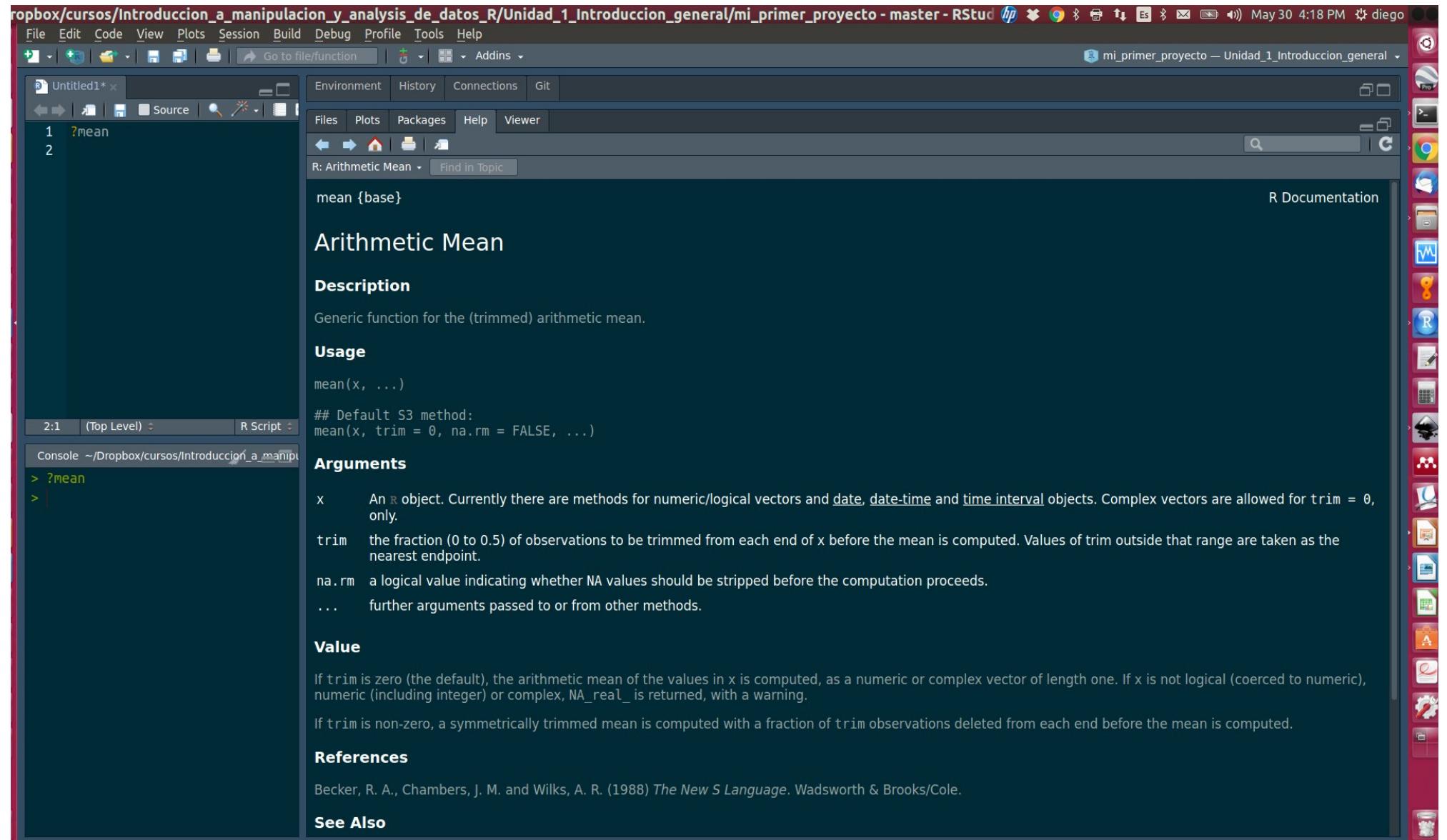
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

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.

> library (lme4)
Loading required package: Matrix
>
```

Cargando el paquete en R

HELPPPPP



Actividad

- Instala R y Rstudio
- Crea en tu sistema de directorios un lugar donde pondrás tus proyectos de R
- Crea un directorio, con el nombre del curso, en donde iras guardando cada Unidad que les envíe.

Tambien aprovecha para instalar en Rstudio varios paquetes que usaremos

Para cargar datos:

```
install.packages ("readxl")
```



Tambien aprovecha para instalar en Rstudio varios paquetes que usaremos

Para cargar datos:

```
install.packages ("readxl")
```

Para manipular datos:

```
install.packages ("tidyverse")  
Install.packages ("stringr")
```

Para visualizar datos:

```
install.packages ("ggplot2")
```

Para análisis de datos:

```
install.packages ("car")  
install.packages ("lme4")  
install.packages ("metafor")  
install.packages ("MCMCglmm")
```

Para presentar datos:

```
install.packages ("markdown")  
install.packages ("xtables")
```

