

Research projects

& programming with R

TropBio

Costa Rica June 4th 2018



Organization for Tropical Studies

where science and nature converge



Ronny



Álvaro

 @alvarovhd

HEREDIA

R User Group



consortium



@RonnyHdezM



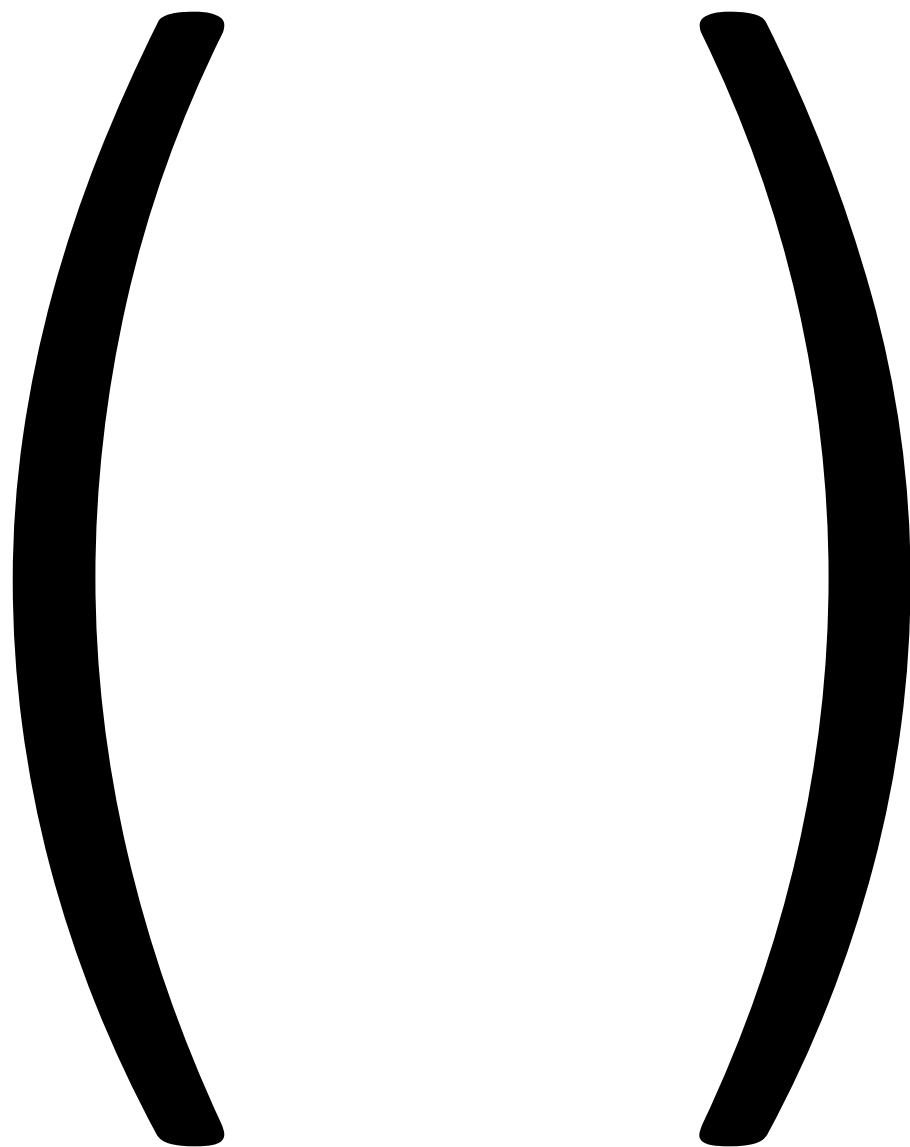
ronnyhdez



ronny.hernandezm@gmail.com

My life with R:



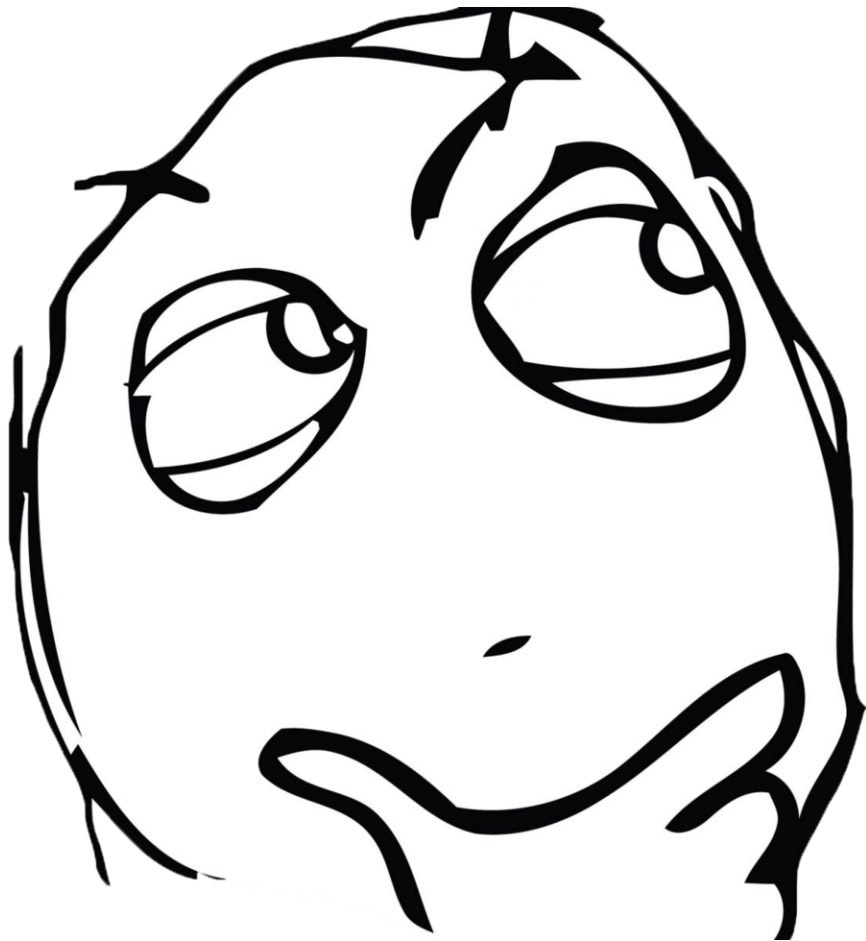


Me: Why do people think I'm mad all the time?

Also me:

I'm not mad! So you can ask me for help



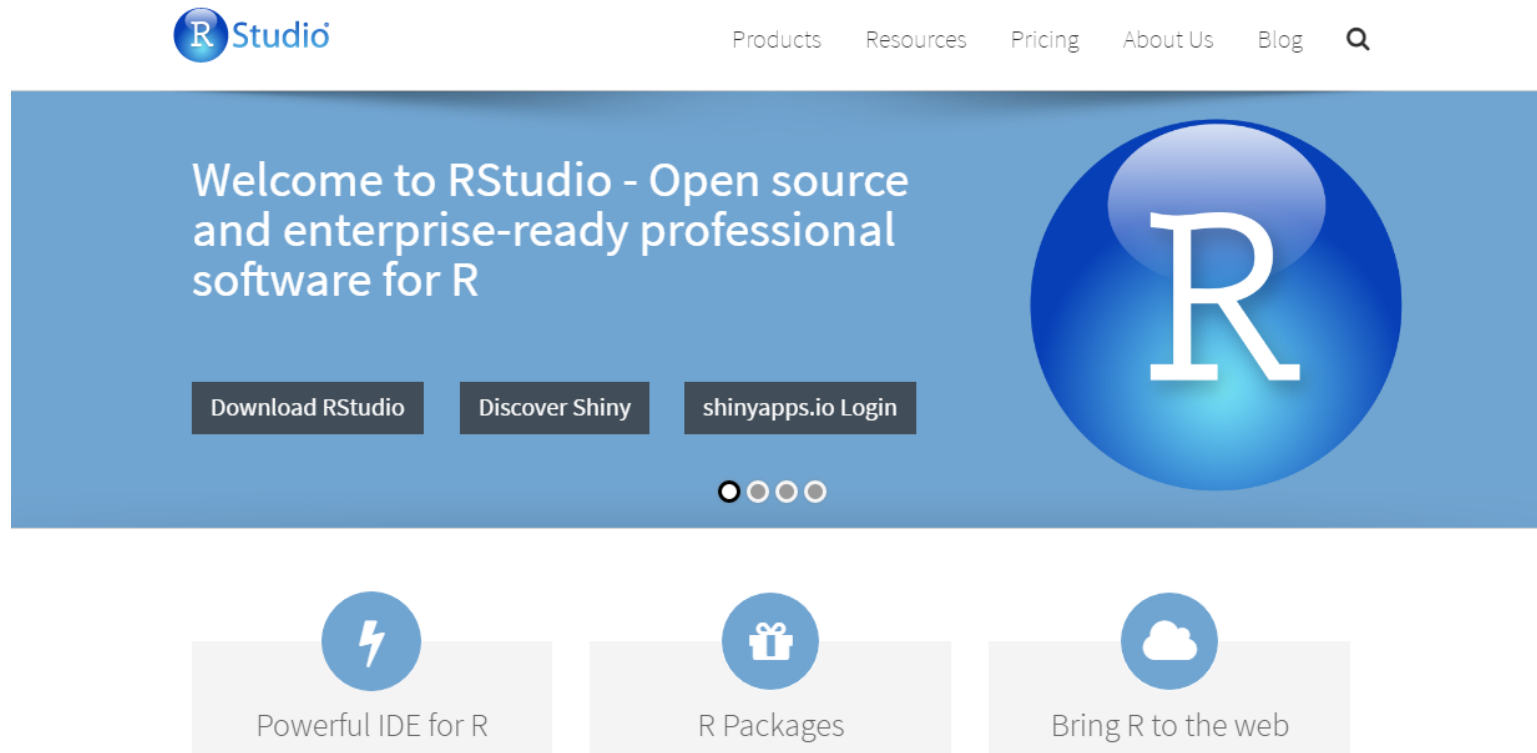


Why R and not a
GUI?

A1   f_x AIRLINE_ID

813444482_T_ONTIME

RStudio






















Integrated Development Environment

Starting a project











Have you ever felt like this
looking at an old analysis?

Bad project structure example:

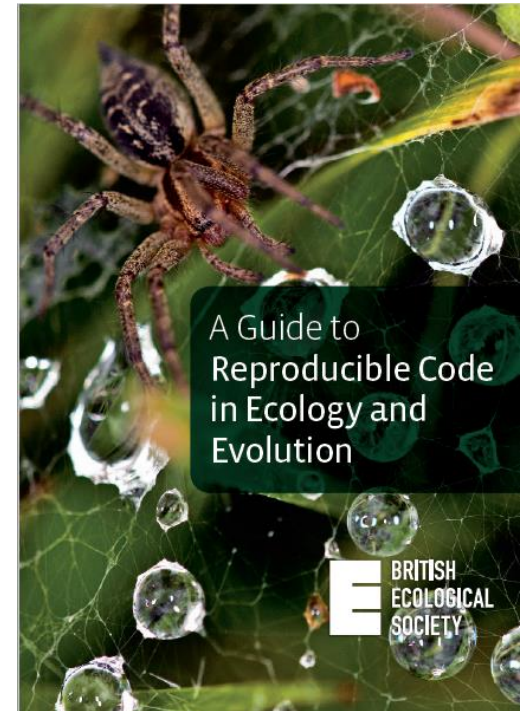
 GAP	05/25/2018 06:43 ...	File folder	
 LAI	05/25/2018 06:43 ...	File folder	
 .Rhistory	10/18/2017 03:52 ...	RHISTORY File	14 KB
 analisis indices canopy.R	11/09/2016 08:27 ...	R File	4 KB
 anotaciones serias.docx	11/24/2016 11:07 ...	Microsoft Word D...	12 KB
 Anotaciones sobre modelos lineales y el t...	10/29/2016 11:04 ...	Microsoft Word D...	14 KB
 data_16.R	10/27/2016 09:16 ...	R File	2 KB
 datastarosa.R	11/10/2016 12:57 ...	R File	7 KB
 Does lianas influence understory species ...	11/18/2016 11:45 ...	Microsoft Word D...	12 KB
 Fotos_b_n_GAP.pptx	11/23/2016 06:46 ...	Microsoft PowerP...	21,032 KB
 Influencia del vecindario y lianas en los r...	11/09/2016 11:09 ...	Microsoft Word D...	401 KB
 Influencia del vecindario y lianas en los r...	11/27/2016 07:01 ...	Microsoft Word D...	376 KB
 Modelos ANCOVA.R	11/10/2016 12:57 ...	R File	2 KB
 NALISIS DE AF VS GROSOR Y PECIOLOS....	10/21/2016 03:27 ...	Microsoft Word D...	27 KB
 nuevo_spinola.R	11/25/2016 09:22 ...	R File	2 KB
 Santa Rosa.pdf	03/27/2017 01:55 ...	PDF File	2,095 KB
 Santa Rosa.pptx	11/10/2016 01:04 ...	Microsoft PowerP...	4,040 KB
 Script_GAP.R	11/24/2016 08:43 ...	R File	3 KB
 script_LAI.R	10/16/2016 02:15 ...	R File	3 KB

Good project structure example:

Name	Date modified	Type	Size
 .git	05/25/2018 06:43 ...	File folder	
 analysis script	05/25/2018 06:43 ...	File folder	
 data	05/25/2018 06:43 ...	File folder	
 doc	05/25/2018 06:43 ...	File folder	
 figs	05/25/2018 06:43 ...	File folder	
 shiny	05/25/2018 06:43 ...	File folder	
 .gitignore	09/12/2017 04:00 ...	GITIGNORE File	1 KB
 README.md	10/03/2017 10:01 ...	MD File	1 KB

Here is an example of a basic project directory structure:

- The **data** folder contains all input data (and metadata) used in the analysis.
- The **doc** folder contains the manuscript.
- The **figs** directory contains figures generated by the analysis.
- The **output** folder contains any type of intermediate or output files (e.g. simulation outputs, models, processed datasets, etc.). You might separate this and also have a **cleaned-data** folder.
- The R directory contains R scripts with function definitions.
- The **reports** folder contains RMarkdown files that document the analysis or report on results.



Why do all this?

1. Reproducibility
2. Automation
3. Communication

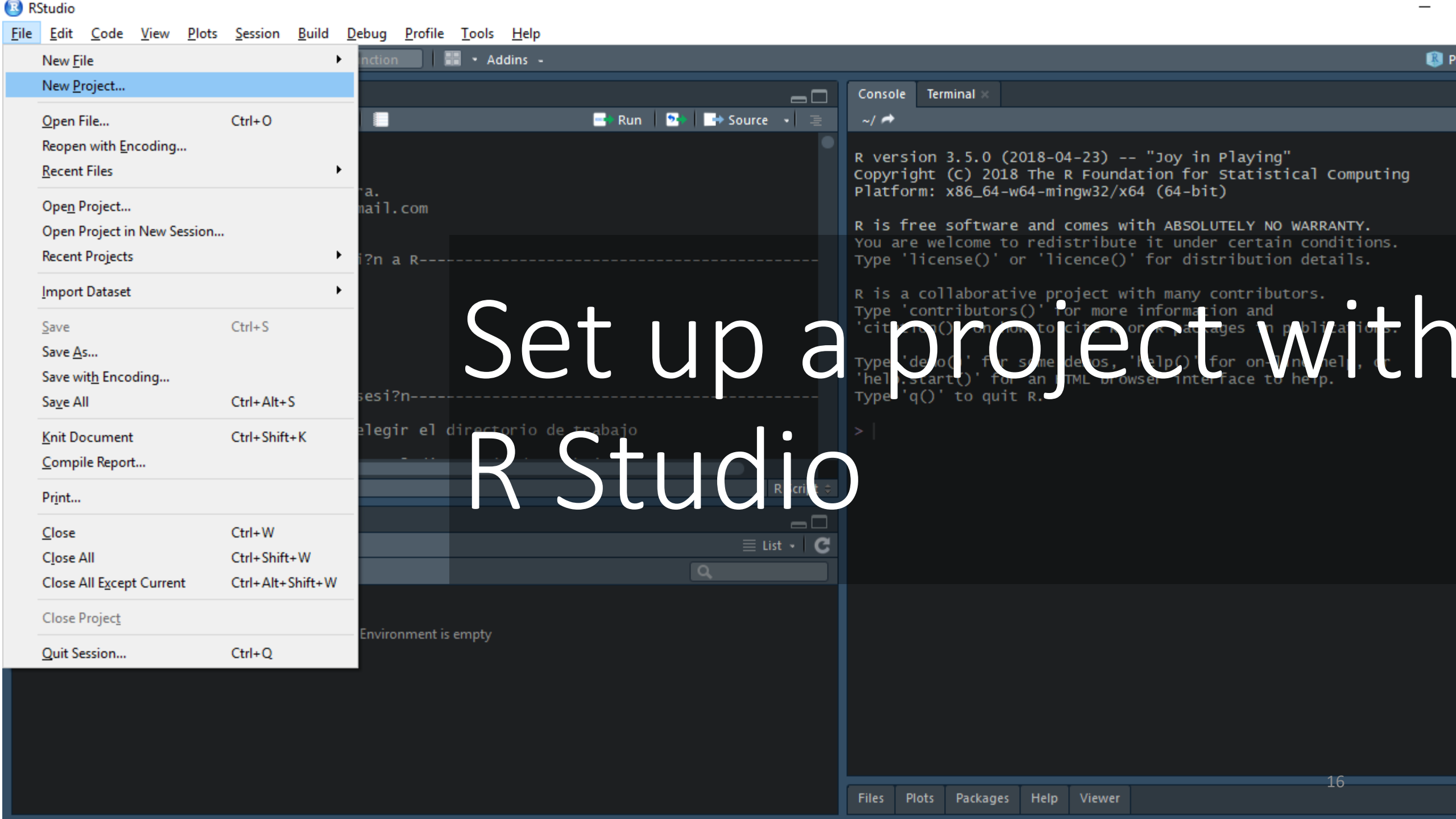


Contribute ↻

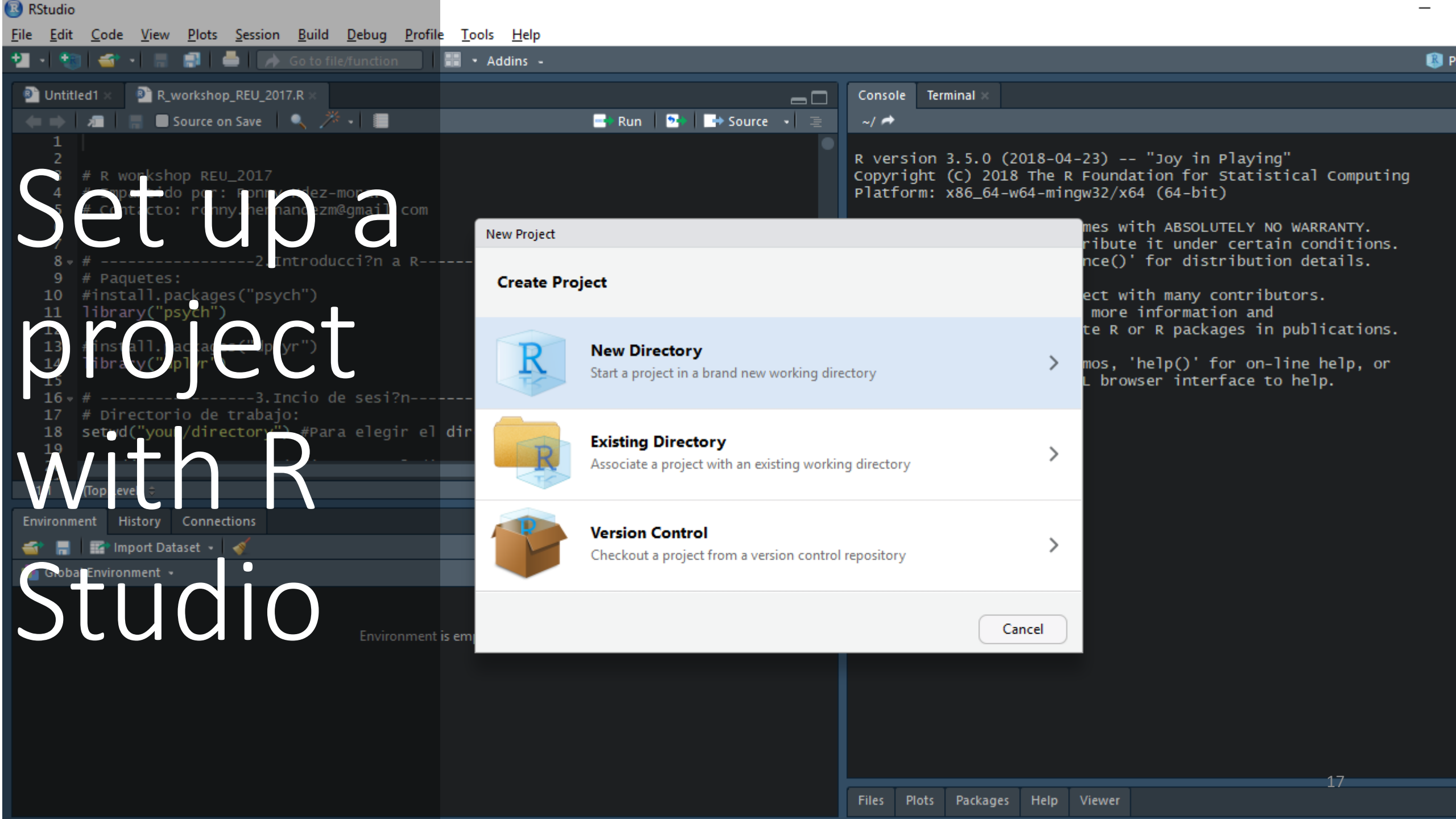
Reproducibility in Science

A Guide to enhancing reproducibility in scientific results and
writing

<http://ropensci.github.io/reproducibility-guide/>



Set up a project with R Studio



Set up a project with RStudio

New Project

Create Project



New Directory

Start a project in a brand new working directory



Existing Directory

Associate a project with an existing working directory



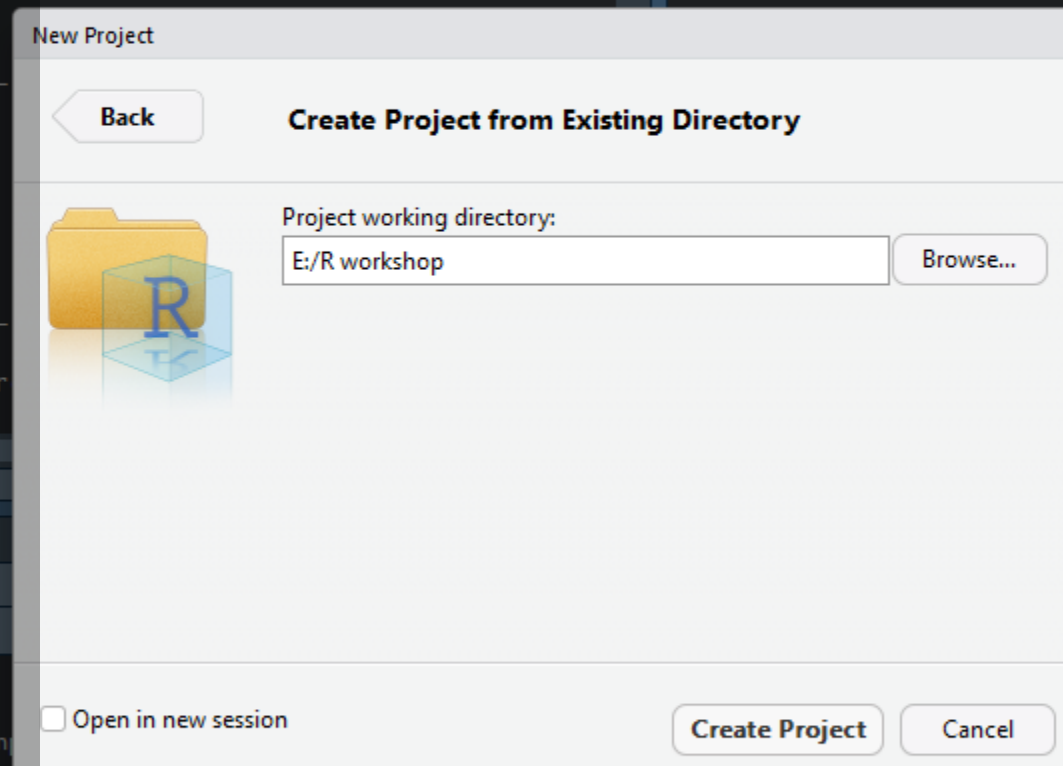
Version Control

Checkout a project from a version control repository



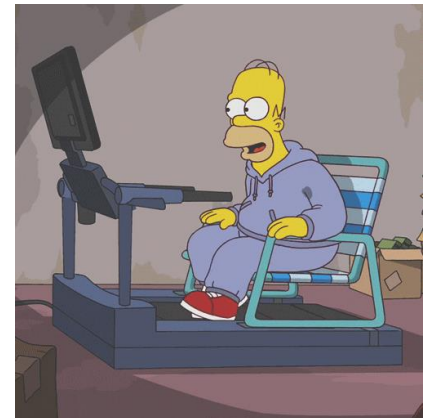
Cancel

Set up a project with R Studio



Exercises

Set the project on Rstudio and the potential structure for your REU research project.



Names:

File names better without spaces: `my_super_pro_data.csv`

Case sensitivity: `My_Super_Pro_data.csv`

Not informative names: `data_1.csv`

Scripts:

Separate scripts: wrangling – analysis - figures

How easy is it to read this?

```
1 eco <- read.csv("~/eco.csv")
2 view(eco)
3 class(eco$Temperatura)
4 cobertura <- subset(eco,Sitio=="cobertura")
5 library(psych)
6 desnudo <- subset(eco,Sitio=="sincobertura")
7 describeBy(desnudo,desnudo$Medici.n)
8 describeBy(cobertura,cobertura$Medici.n,na.rm=T)
9 describeBy(eco,eco$Sitio,na.rm=T)
10 sitio <- c("cobertura","cobertura","cobertura","sin","sin","sin")
11 lux <- c(8800,13300,12800,71000,104600,118400)
12 iluminancia <- data.frame(sitio,lux)
13 humedad <- c(37.1,37.9,38.7,39.9,39,36.8)
14 hume <- data.frame(sitio,humedad)
15 iluminancia
16 class(iluminancia$lux)
17 describeBy(iluminancia,iluminancia$sitio)
18 describeBy(hume,hume$sitio)
19 z <- read.table(file="clipboard",sep="\t",header=T)
20 plot(x=z$?rea.acumulada,y=z$Spp.acumulada,pch=16,xlab=expression(paste("?rea(",cm^2,")",sep="")),ylab="Especies acumuladas")
21 lines.default(z$?rea.acumulada,z$Spp.acumulada)
```

1:1 (Top Level) ↕

R Script ↕

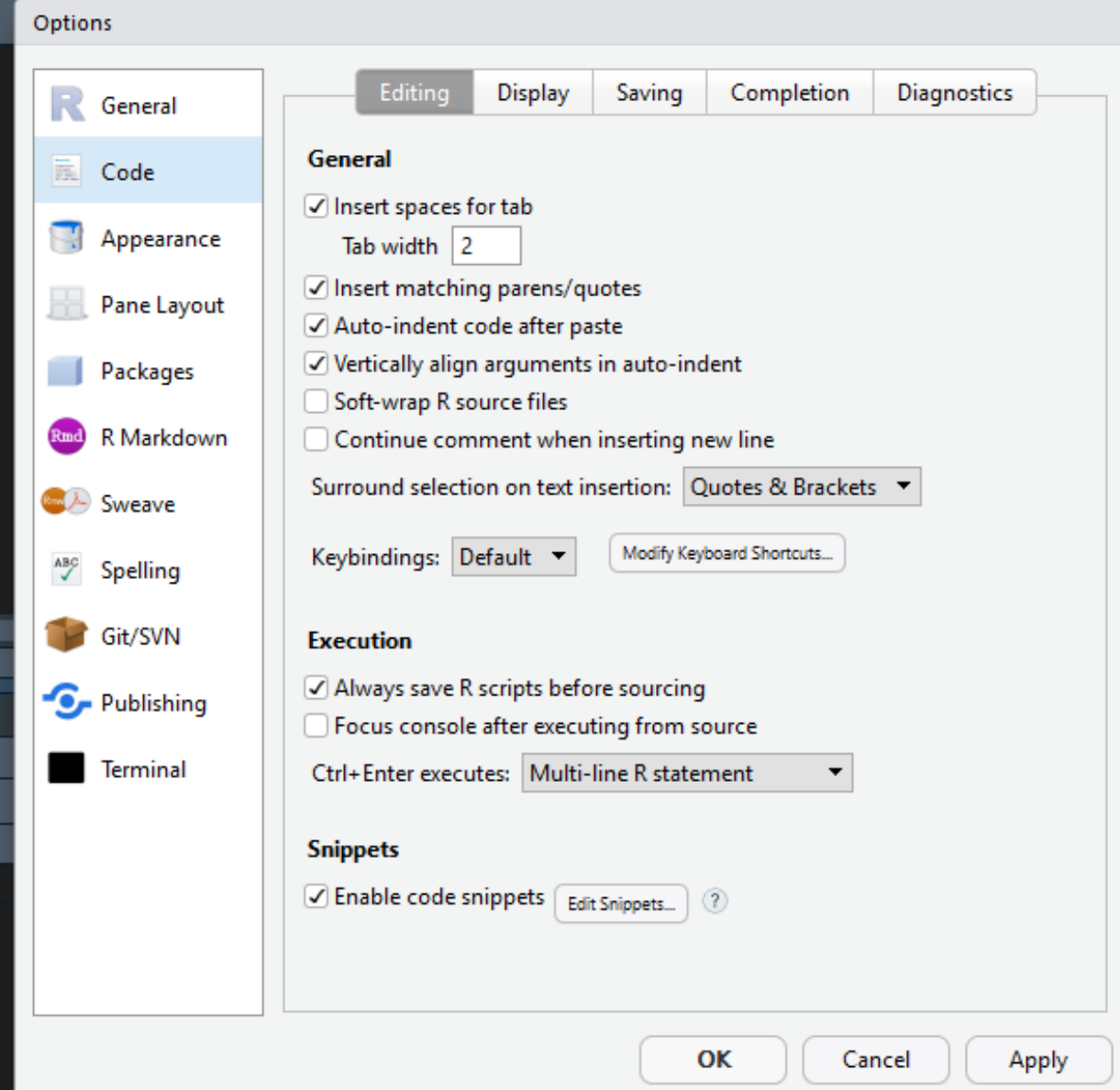
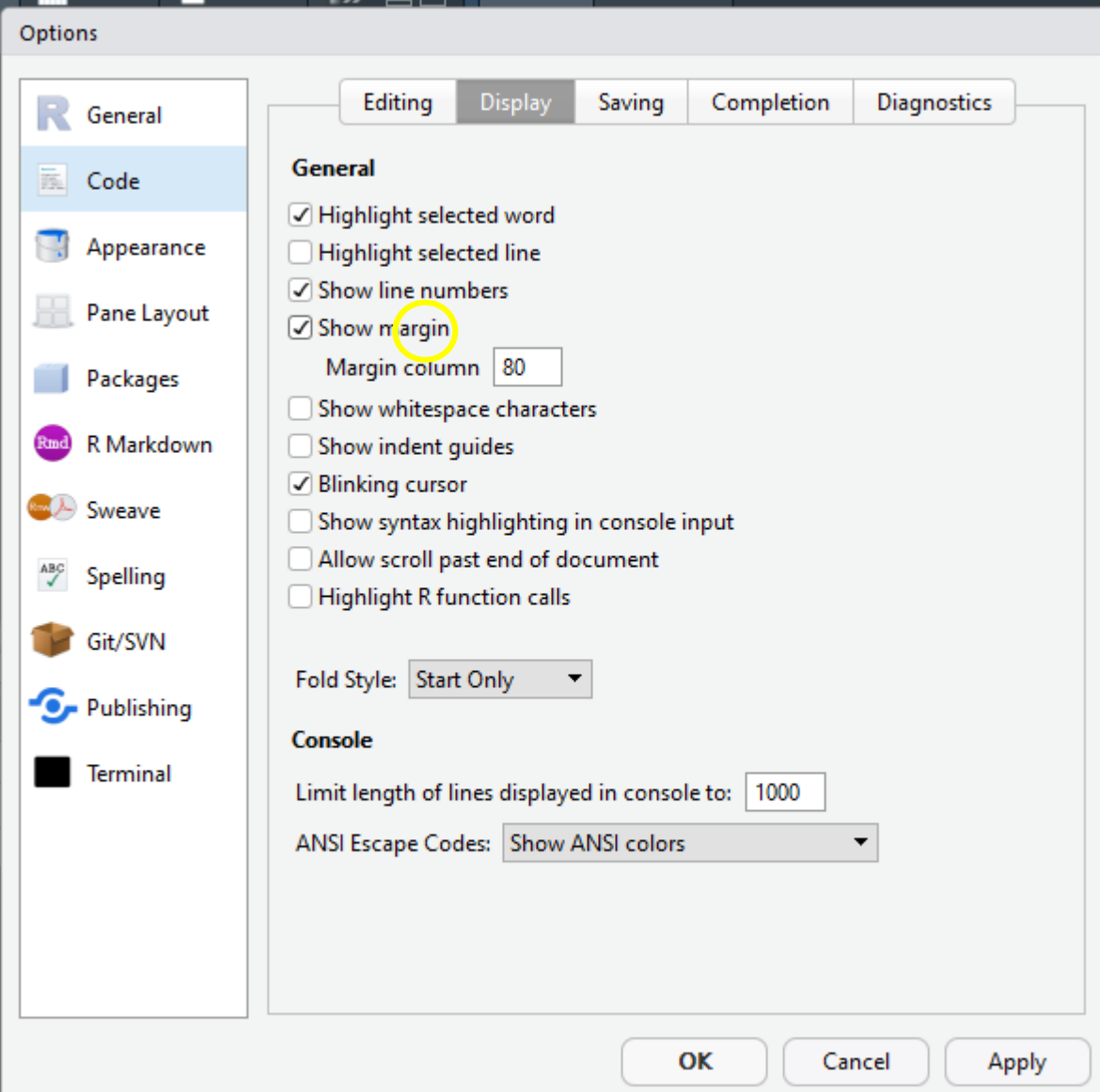
How easy is it to read this?

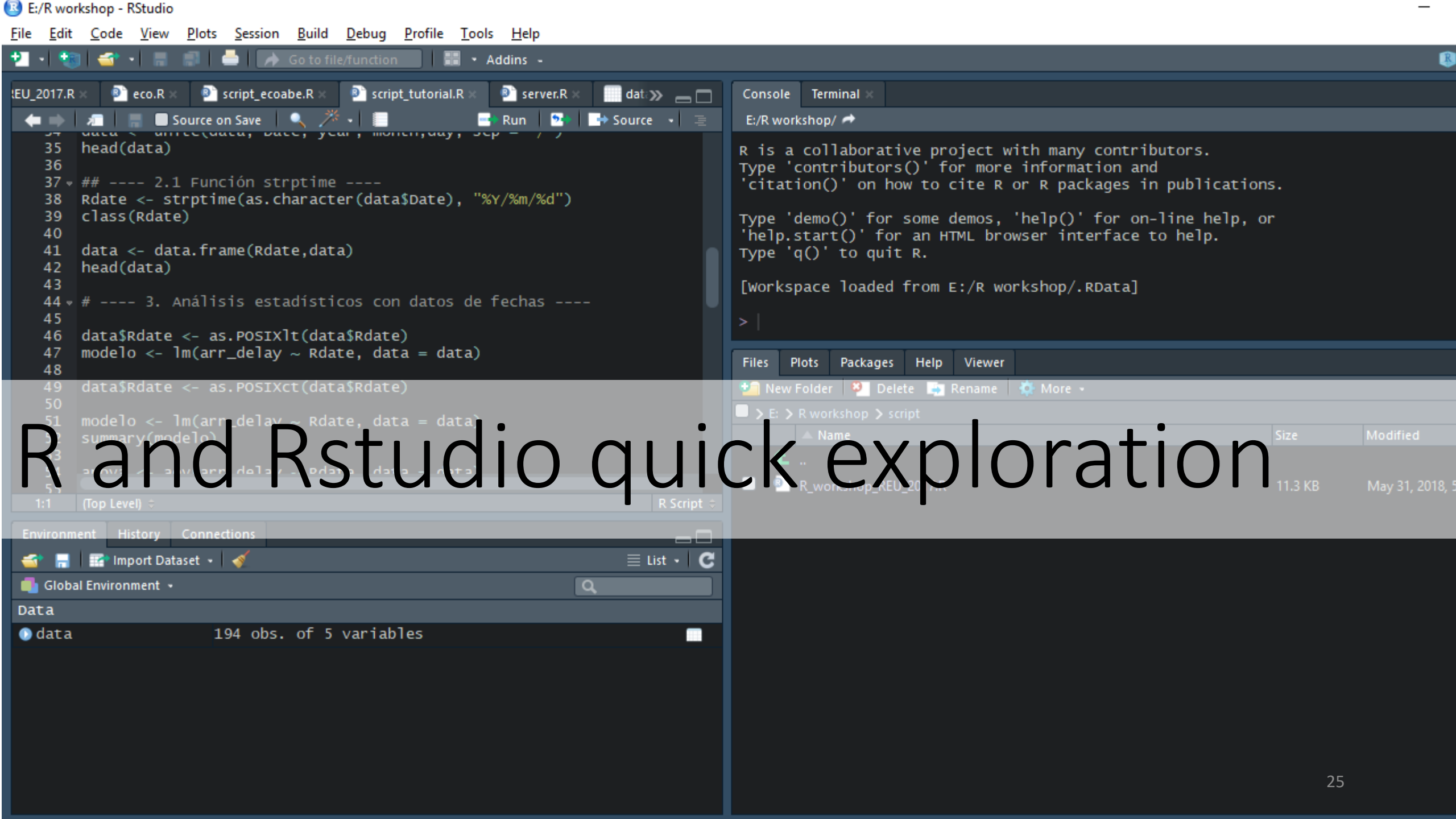
```
31 #-----Con paquete vegan-----  
32 #Primero cargo los datos por sitio, no por d?a para sacar ?ndices por sitio  
33  
34 site <- read.csv("site.csv",header=T,row.names=1)  
35 diversity(persite,index = "shannon")  
36 diversity(persite,index="simpson")  
37  
38 #Margalef:  
39 n <- apply(persite>0,1,sum)  
40 N <- apply(persite,1,sum)  
41 (n-1)/log(N)  
42  
43  
44 #Luego debo de hacer curva de acumulaci?n por sitio, por lo tanto  
45 #debo de hacer el subset respectivo  
46  
47 una <- subset(site,Sitio=="UNA")  
48 tail(una)  
49 una <- una[1:7,-1]  
50  
51 sta_lu <- subset(site,Sitio=="Sta_Luc?a")  
52 tail(sta_lu)  
53 sta_lu <- sta_lu[,-1]  
54  
55 monte <- subset(site,Sitio=="Monte_Cruz")  
56 tail(monte)  
57 monte <- monte[,-1]  
58  
59 #Para curvaaccu debo de cargar por muestra (por d?a)  
60  
61 spa <- specaccum(una,method="random")  
62 spa  
63 plot(spa,xlab = "D?as de Muestreo",ylab = "Riqueza",col="red")  
64  
65 spasta <- specaccum(sta_lu)  
66 plot(spasta,xlab = "D?as de Muestreo",ylab = "Riqueza",add=T)
```

How easy is it to read this?

```
34 data <- unlist(data, date, year, month, day, sep = "/")
35 head(data)
36
37 ## ---- 2.1 Función strptime ----
38 Rdate <- strptime(as.character(data$Date), "%Y/%m/%d")
39 class(Rdate)
40
41 data <- data.frame(Rdate, data)
42 head(data)
43
44 # ---- 3. Análisis estadísticos con datos de fechas ----
45
46 data$Rdate <- as.POSIXlt(data$Rdate)
47 modelo <- lm(arr_delay ~ Rdate, data = data)
48
49 data$Rdate <- as.POSIXct(data$Rdate)
50
51 modelo <- lm(arr_delay ~ Rdate, data = data)
52 summary(modelo)
53
54 anova <- aov(arr_delay ~ Rdate, data = data)
55 summary(anova)
56
57 #convertir a continuo
58 data$Rdate <- as.POSIXct(data$Rdate)
59
60 ## ---- 3.1 Nombre completo del día ----
61 ejemplo <- data[1,1]
62 ejemplo
63
64 weekdays(ejemplo)
65
66 ## ---- 3.2 Otros formatos ----
67 otras_fechas <- c("2feb2016", "18jun1990", "7nov1995")
68 strptime(otras_fechas, "%d%b%Y")
69
70 # ---- 4. cálculos con el tiempo ----
```

1:1 (Top Level) R Script







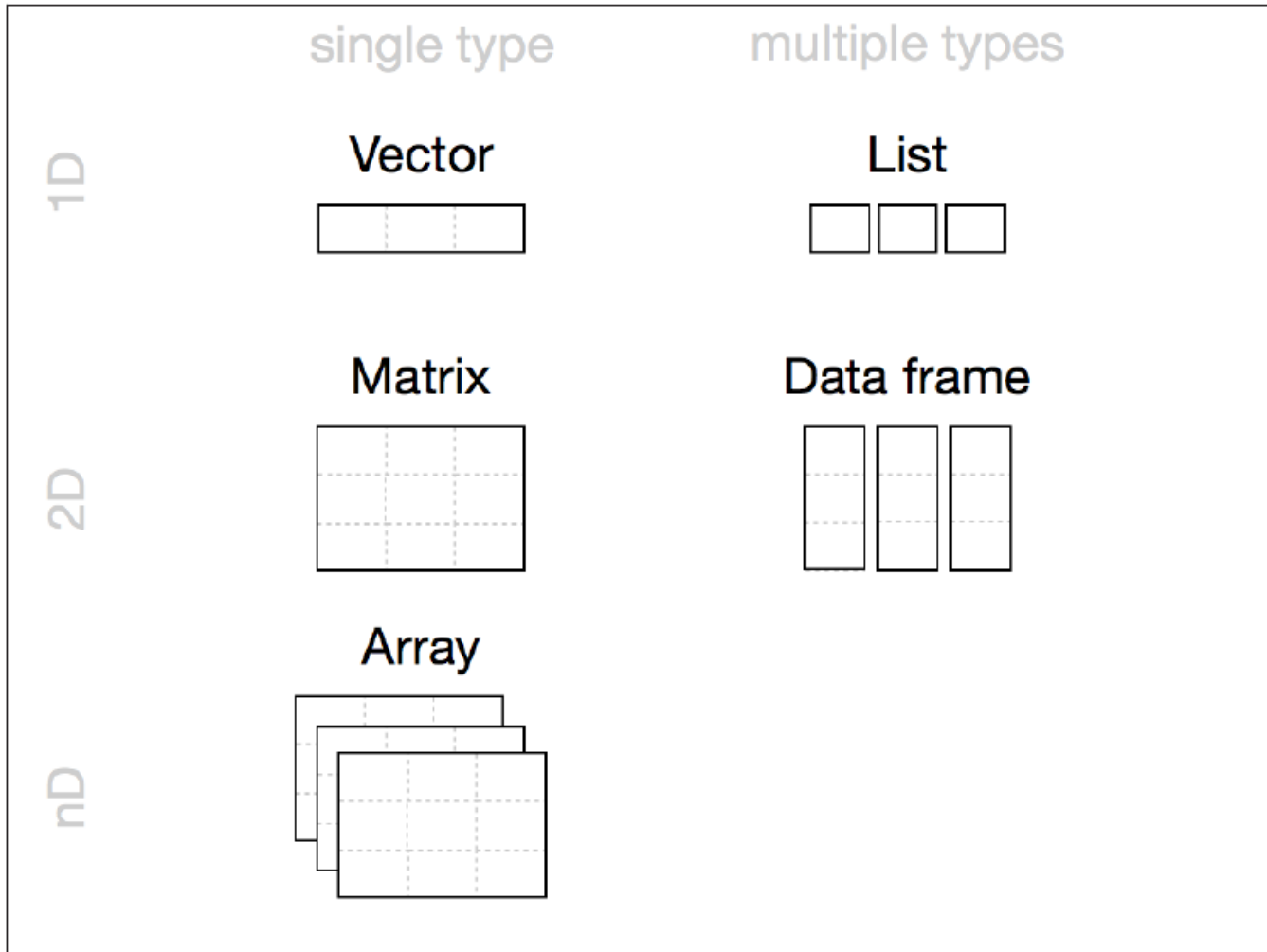
Data and structures in R

Types of data:

-Characters	“biología” , “A” , “estadística”
-Numerics	18.6 , 90
-Integers	18L , 26L
-Complex	2+4i
-Logical	TRUE , FALSE

Para verificar su clase: `class()`

Objetos se almacenan en estructuras:



>Check_the_code::



Coercing rules

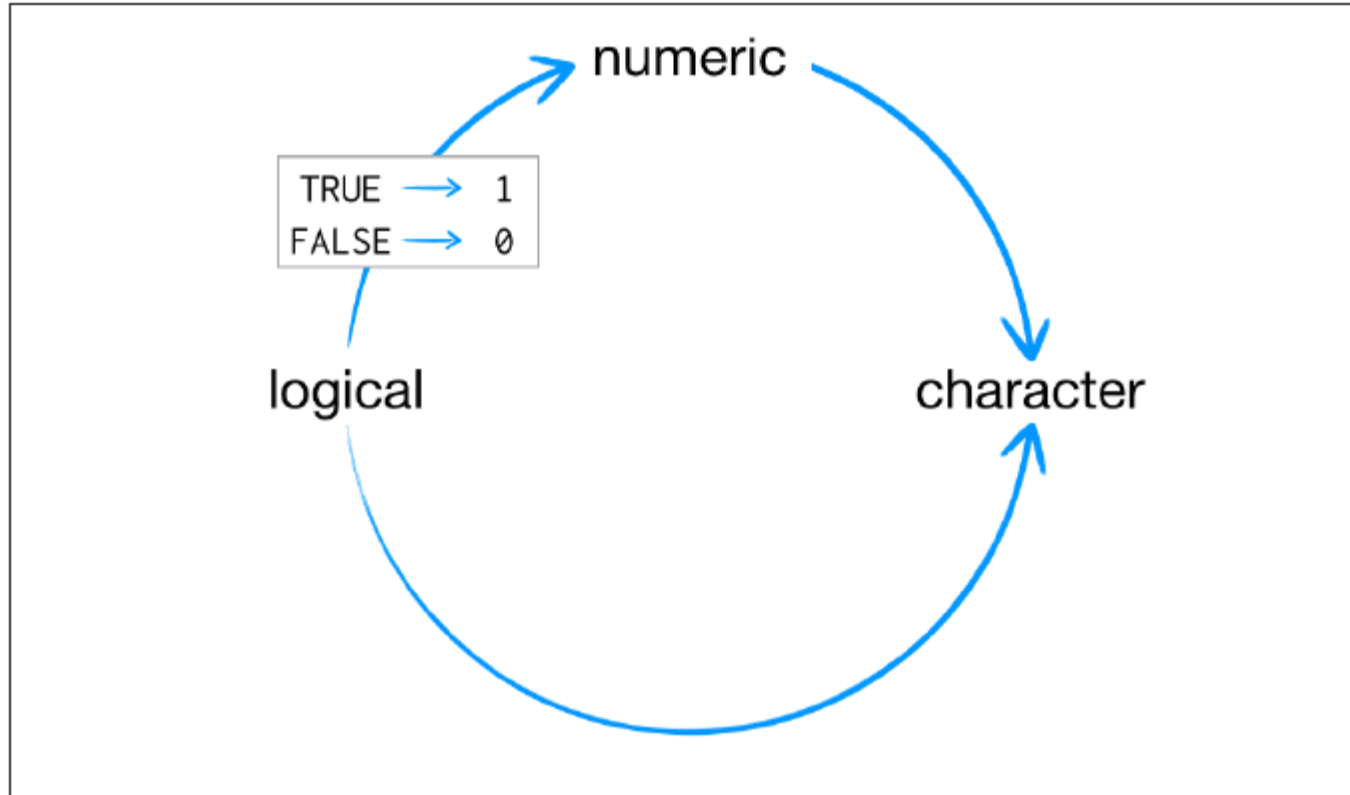


Figure 3-1. R always uses the same rules to coerce data to a single type. If character strings are present, everything will be coerced to a character string. Otherwise, logicals are coerced to numerics.

>Check_the_code::



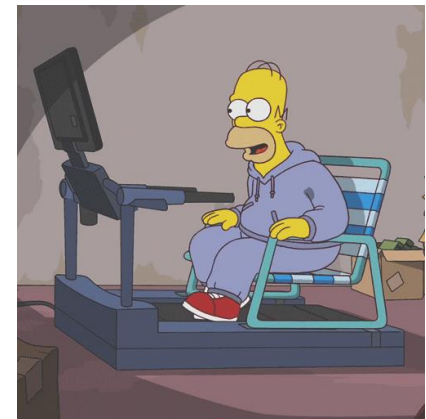
Exercises

1. Why this is not a numeric class?

```
numeric <- ('1:1000')
```

2. Obtain the **mean** of the **Aniso.grey** column from the **Camara_3** dataset.

Hint: `mean(dataname$columnname)`



A small, light-colored insect is shown on a light-colored, textured surface, juggling five colorful balls. The balls are in various positions: one is in the air above the insect, another is on its back, and three others are in the air above that. The balls are red, blue, and multi-colored. The background is dark and out of focus.

Loops and functions basics

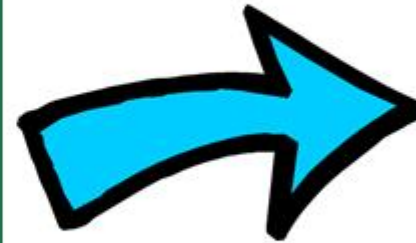
Reduce duplication!

- Code easier to read
- Easier to respond to changes in requirements
- Fewer bugs





Data wrangling



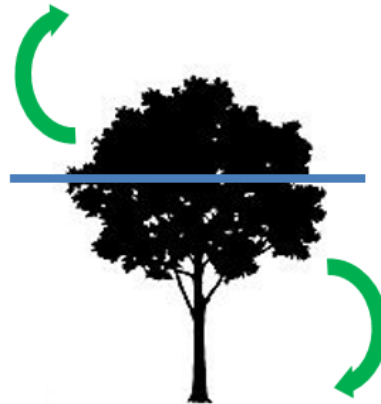
Don't mix characters and numbers in the same cell
Put easy names to write and remember
Don't mix lower case and upper case

Experimental design:

- 3 Individuals
- 2 measurement per leaf
- 5 treatments applied to 2 strata

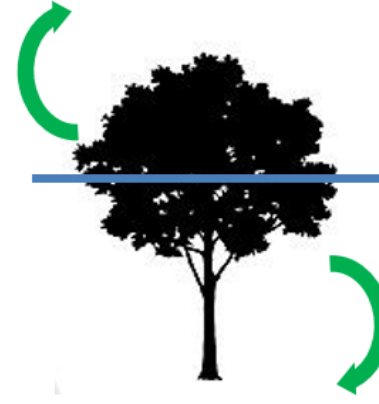
How should the data set be structured?

10 HD	27 °C
10 HD	34°C
10 HD	38 °C
10 HD	45 °C
10 HD	EH



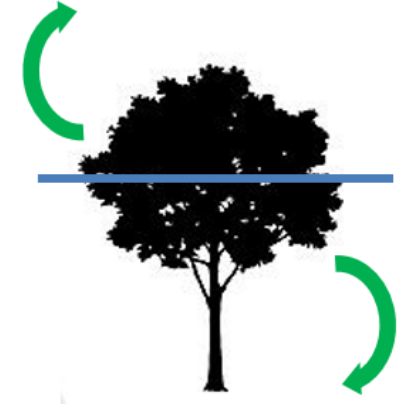
10 HS	27 °C
10 HS	34°C
10 HS	38 °C
10 HS	45 °C
10 HS	EH

10 HD	27 °C
10 HD	34°C
10 HD	38 °C
10 HD	45 °C
10 HD	EH



10 HS	27 °C
10 HS	34°C
10 HS	38 °C
10 HS	45 °C
10 HS	EH

10 HD	27 °C
10 HD	34°C
10 HD	38 °C
10 HD	45 °C
10 HD	EH



10 HS	27 °C
10 HS	34°C
10 HS	38 °C
10 HS	45 °C
10 HS	EH

DATA_FIXb - Microsoft Excel

ARCHIVO INICIO INSERTAR DISEÑO DE PÁGINA FÓRMULAS DATOS REVISAR VISTA

Pegar Fuente Alineación Número Estilos Celdas Modificar

Ordenar y filtrar Buscar y seleccionar Fírmalo y Codifíquelo

Inicio sesión

I19

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Inicial					Final										
2	ID	IND	POS	TRT		ID	IND	POS	TRT							
3		1 A	D	27		170 B	S		34							
4		2 A	D	27		261 C	S		34							
5		3 A	D	27		262 C	S		34							
6		4 A	D	27		263 C	S		34							
7		5 A	D	27		264 C	S		34							
8		6 A	D	27		265 C	S		34							
9		7 A	D	27		266 C	S		34							
10		8 A	D	27		267 C	S		34							
11		9 A	D	27		268 C	S		34							
12		10 A	D	27		269 C	S		34							
13		101 B	D	27		270 C	S		34							
14		102 B	D	27		61 A	S		34							
15		103 B	D	27												
16		104 B	D	27												
17		105 B	D	27												
18		106 B	D	27												
19		107 B	D	27												
20																
21																
22																
23																

DATA_FIXb Hoja1

LISTO

100%

Microsoft Excel interface showing a spreadsheet titled "DATA_FIXb". The ribbon includes tabs: ARCHIVO, INICIO, INSERTAR, DISEÑO DE PÁGINA, FÓRMULAS, DATOS, REVISAR, and VISTA. The "INICIO" tab is active, displaying options for Font (Fuente), Alignment (Alineación), Number (Número), Styles (Estilos), Cells (Celdas), and Modify (Modificar).

The spreadsheet data is as follows:

Inicial			
ID	IND	POS	TRT
1	A	D	27
2	A	D	27
3	A	D	27
4	A	D	27
5	A	D	27
6	A	D	27
7	A	D	27
8	A	D	27
9	A	D	27
10	A	D	27
101	B	D	27
102	B	D	27
103	B	D	27
104	B	D	27
105	B	D	27
106	B	D	27
107	B	D	27
Final			
ID	IND	POS	TRT
170	B	S	34
261	C	S	34

A large red prohibition sign (a circle with a diagonal line) is overlaid on the right side of the spreadsheet, covering columns J through P. The sign is centered vertically and horizontally over the data area.

DATA_FIXb - Microsoft Excel

ARCHIVO INICIO INSERTAR DISEÑO DE PÁGINA FÓRMULAS DATOS REVISAR VISTA

Portapapeles Fuente Alineación Número Estilos Celdas Modificar Privacidad

A1 Final

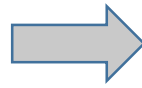
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Final															
2	IND	B	C	A												
3	POS	S	D	S												
4	TRT	34	34	34												
5	FV	X	X	X												
6																
7																
8																
9																
10																
11																
12																
13																
14																
15																
16																
17																
18																
19																
20																
21																
22																
23																
24																

DATA_FIXb Hoja1

LISTO 100%

Tidy data!

ID	IND	POS	TRT	Medicion	FV
1	A	D	27	F	
2	A	S	27	F	
3	B	D	34	I	
4	B	S	34	I	
5	C	D	38	F	



Observation



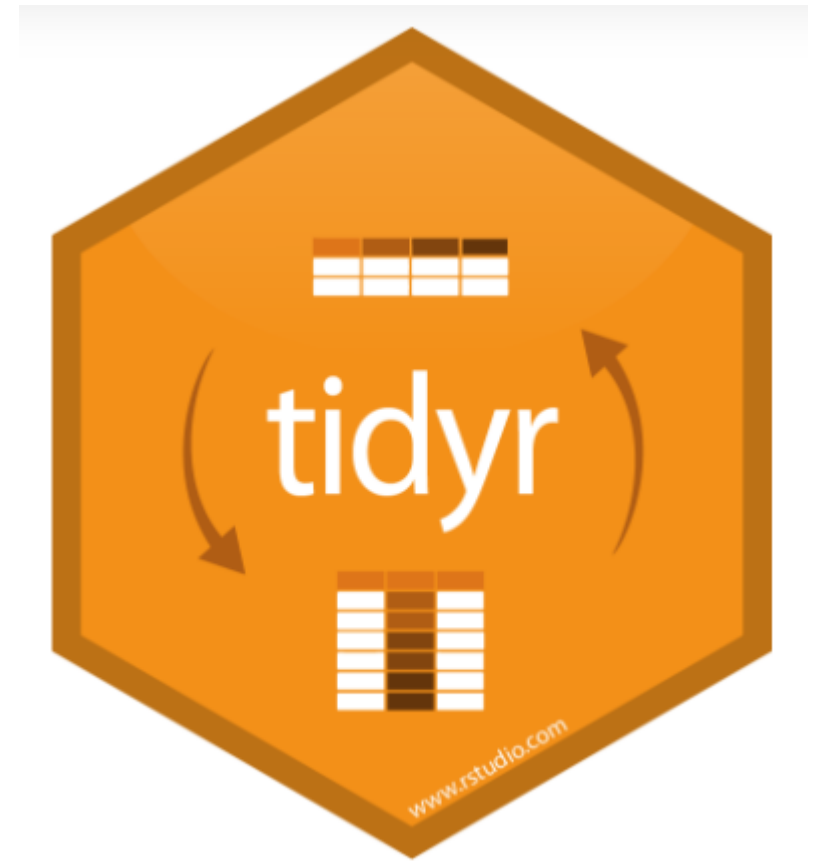
Variable / Attribute

- Each variable must have its own column.
- Each observation must have its own row.
- Each value must have its own cell.

Wickham, H. (2014). Tidy Data. *Journal of Statistical Software*, 59(i10).

Four verbs in tidyr:

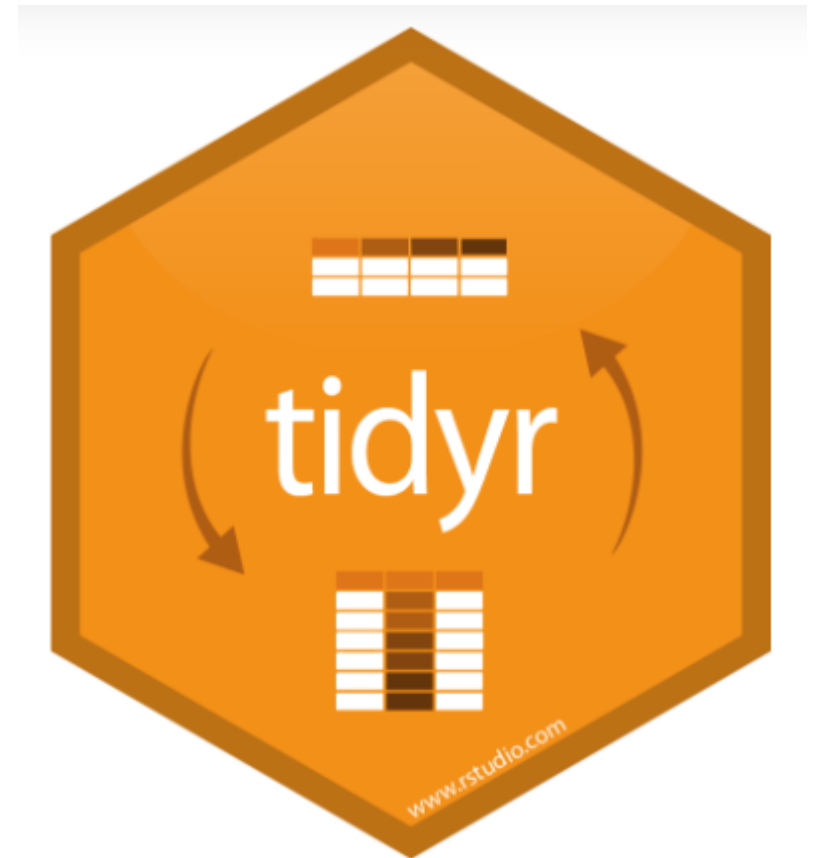
- 1- `gather ()`
- 2- `spread ()`
- 3- `separate()`
- 4- `unite ()`



How to deal with not tidy data?

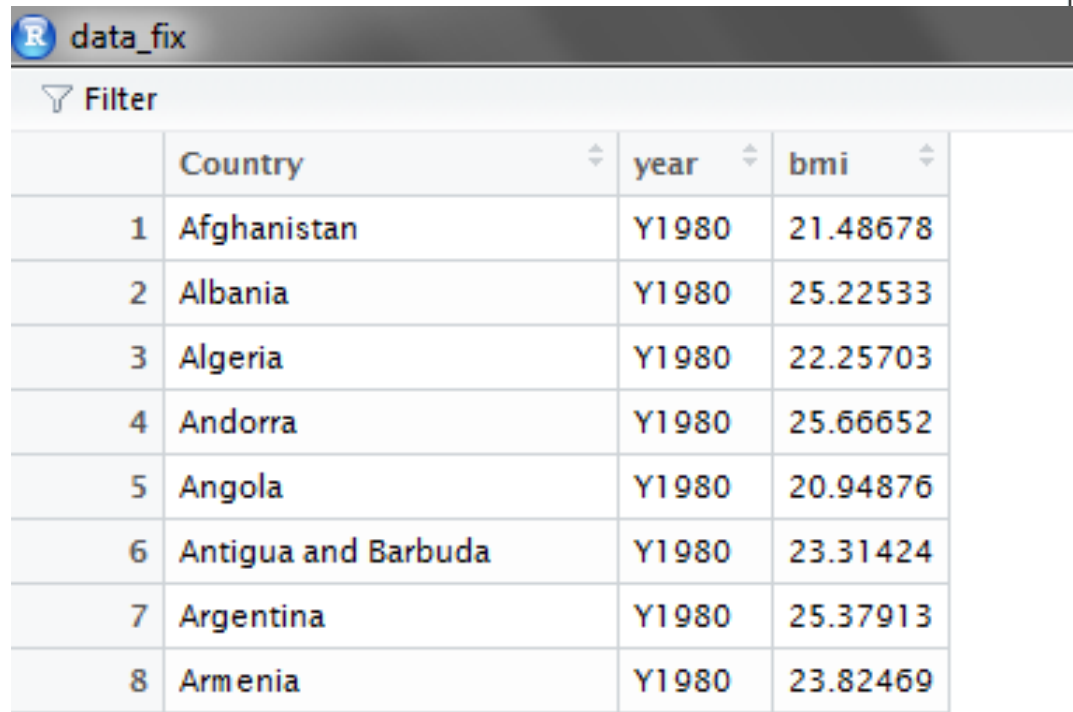
What is the problem with this dataset?

R data					
Filter					
	Country	Y1980	Y1981	Y1982	Y1983
1	Afghanistan	21.48678	21.46552	21.45145	21.43822
2	Albania	25.22533	25.23981	25.25636	25.27176
3	Algeria	22.25703	22.34745	22.43647	22.52105
4	Andorra	25.66652	25.70868	25.74681	25.78250
5	Angola	20.94876	20.94371	20.93754	20.93187
6	Antigua and Barbuda	23.31424	23.39054	23.45883	23.53735
7	Argentina	25.37913	25.44951	25.50242	25.55644
8	Armenia	23.82469	23.86401	23.91023	23.95649
9	Australia	24.92729	25.00216	25.07660	25.14938
10	Austria	24.84097	24.88110	24.93482	24.98118
11	Azerbaijan	24.49375	24.52584	24.56064	24.60150



How to deal with not tidy data?

What is the problem with this set of data?



	Country	year	bmi
1	Afghanistan	Y1980	21.48678
2	Albania	Y1980	25.22533
3	Algeria	Y1980	22.25703
4	Andorra	Y1980	25.66652
5	Angola	Y1980	20.94876
6	Antigua and Barbuda	Y1980	23.31424
7	Argentina	Y1980	25.37913
8	Armenia	Y1980	23.82469

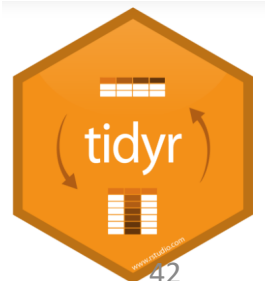
```
gather(data, key, value, ...)
```

data: a data frame

key: bare name of new key column

value: bare name of new value column

..: bare names of columns to gather (or not)



How to deal with not tidy data?

R data_unfix

Filter

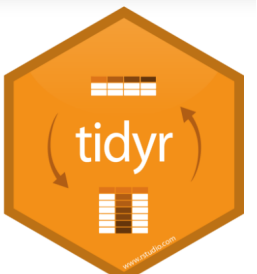
	Country	Y1980	Y1981	Y1982	Y1983
1	Afghanistan	21.48678	21.46552	21.45145	21.43822
2	Albania	25.22533	25.23981	25.25636	25.27176
3	Algeria	22.25703	22.34745	22.43647	22.52105
4	Andorra	25.66652	25.70868	25.74681	25.78250
5	Angola	20.94876	20.94371	20.93754	20.93187
6	Antigua and Barbuda	23.31424	23.39054	23.45883	23.53735
7	Argentina	25.37913	25.44951	25.50242	25.55644
8	Armenia	23.82469	23.86401	23.91023	23.95649
9	Australia	24.02720	25.00216	25.07660	25.14038

spread(data, key, value)

data: a data frame

key: bare name of column containing keys

value: bare name of column containing values



How to deal with not tidy data?

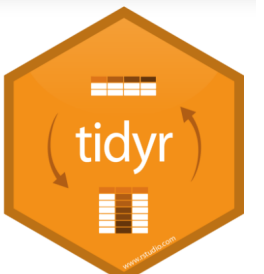
How to solve this?

R fotos_sep

Filter

	i	plot	MIG.grey	Aniso.grey	Timestamp	Date	Time	sitio	camara
1	1	LADERA	0.2220831	0.9363847	21/06/2016 11:47	2015:05:21	13:21:50	Piro	1
2	2	LADERA	0.2334211	0.9371384	21/06/2016 11:47	2015:05:21	14:21:50	Piro	1
3	3	LADERA	0.1685896	0.9400231	21/06/2016 11:48	2015:05:22	09:00:01	Piro	1
4	4	LADERA	0.1554022	0.9261470	21/06/2016 11:48	2015:05:22	10:00:01	Piro	1
5	5	LADERA	0.2006901	0.9033331	21/06/2016 11:48	2015:05:22	11:00:01	Piro	1

>Check_the_code::

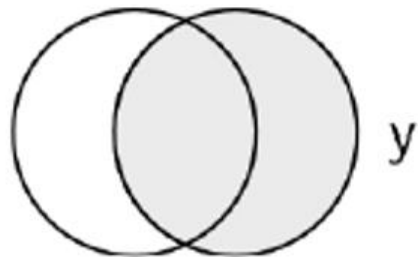
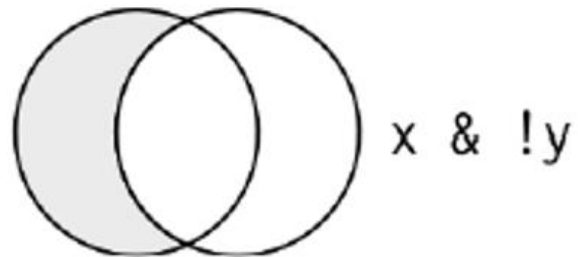
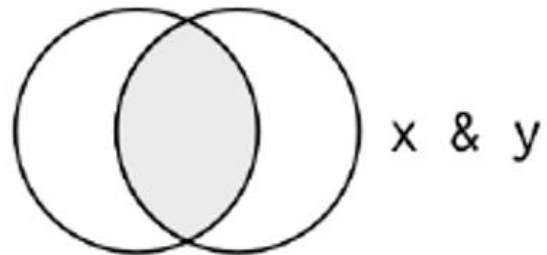
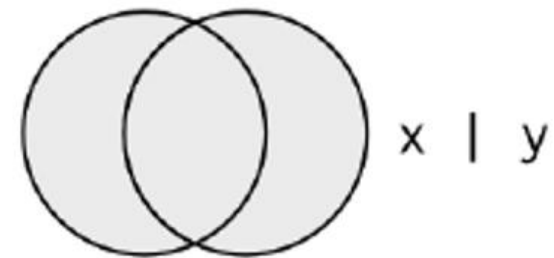
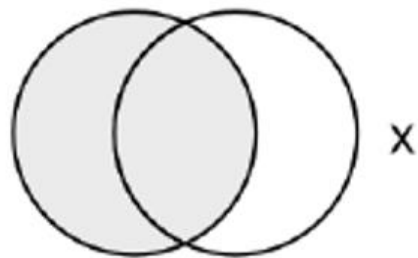
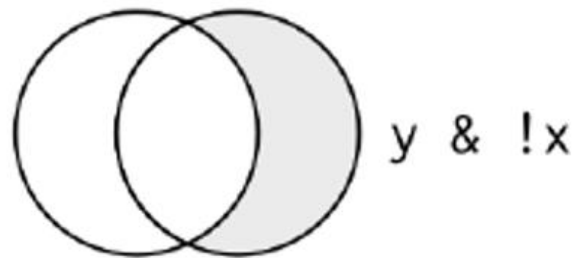


Five verbs in dplyr:

- 1- `select ()`
- 2- `filter ()`
- 3- `arrange ()`
- 4- `mutate ()`
- 5- `summarise ()`
 - *- `group_by ()`



Logical operators:



>Check_the_code::



ggplot2:

```
ggplot(data = <DATA>) +  
  <GEOM_FUNCTION> (  
    mapping = aes(<MAPPINGS>),  
    stat = <STAT>,  
    position = <POSITION>  
  ) +  
  <COORDINATE_FUNCTION> +  
  <FACET_FUNCTION>
```

>Check_the_code::



Save R objects for the next session:

```
save(my_data, file = "my_data.Rdata")
```

```
saveRDS(mtcars, "my_data.rds")
```

```
readRDS(file = "my_data.rds")
```


Repositories:

- CRAN, <http://cran.r-project.org>
- Bitbucket, <https://bitbucket.org>
- Bioconductor, <http://www.bioconductor.org>
- GitHub, <https://github.com>
- Gitorious, <http://www.gitorious.com>

Help:

The screenshot displays the RStudio environment with three main panels:

- Script Editor (Left):** Contains R code for data manipulation and outlier detection. The code includes comments in Spanish and uses functions like `subset`, `order`, `head`, `tail`, and `plot`.
- Console (Top Right):** Shows the execution of R commands. It includes the creation of random normal distributions for three semesters and the calculation of their means using `lapply` and `sapply`.
- Help Window (Bottom Right):** The **Help** menu item in the top toolbar is circled in yellow. The help window displays the documentation for the `strptime` function, titled "Date-time Conversion Functions to and from Character".

```
328 #Diapo 93
329 #Otros
330 #Para ordenar de menor a mayor:
331 orden <- fotos_sep[order(fotos_sep$MIG.grey),]
332 head(orden)
333 tail(orden)
334
335 #O bien, ordenamos subconjuntos con su valor:
336 orden_1 <- fotos_sep[order(fotos_sep$plot,fotos_sep$MIG.grey),]
337 head(orden_1,20)
338 tail(orden_1,20)
339
340 #Diapo 96
341 #Subset haciendo uso de operadores:
342 #Estos son útiles para indicar a R ciertas
343 #características que deseamos seleccionar:
344
345 outliers_1 <- subset(fotos_sep,select=plot,subset=(MIG.grey>0.4),)
346 outliers_2 <- fotos_sep[fotos_sep$MIG.grey>0.4,]
347 head(outliers)
348
349 #Recordemos del apartado de las fechas, que en el plot
350 #detectamos unos outliers. Con subsets y operadores
351 #lógicos los podemos eliminar
352 eliminate <- subset(fotos_sep,subset=MIG.grey<0.33)
353 head(eliminate)
354 plot(eliminate$fecha,eliminate$MIG.grey)
355
```

```
> x <- list(I_SEMESTRE=rnorm(10,8,1),
+           II_SEMESTRE=rnorm(10,7,1),
+           III_SEMESTRE=rnorm(10,7,1))
> lapply(x,mean)
$I_SEMESTRE
[1] 7.934191

$II_SEMESTRE
[1] 6.94863

$III_SEMESTRE
[1] 6.770249

> sapply(x,mean)
  I_SEMESTRE  II_SEMESTRE  III_SEMESTRE 
7.934191    6.948630    6.770249
```

Help Window Content:

strptime (base)

Date-time Conversion Functions to and from Character

Description

Functions to convert between character representations and objects of classes "POSIXlt" and "POSIXct" representing calendar dates and times.

Usage

Help:

The screenshot shows the Stack Overflow homepage. At the top is a dark navigation bar with the StackExchange logo, user avatars, and a search bar. Below this is the main header with the Stack Overflow logo and navigation links: Questions, Jobs, Documentation (Beta), Tags, Users, Badges, and Ask Question. A large blue banner announces 'Stack Overflow Documentation' with the text 'We started with Q&A. Technical documentation is next, and we need your help. Whether you're a beginner or an experienced developer, you can contribute.' and a button 'I want to help →'. To the right of the banner is a preview of the documentation interface. Below the banner, a question is displayed: 'How to subset a data frame, apply a function with a for-loop and obtain the result with the name of each subset on it?'. The question is marked as 'asked 1 year ago', 'viewed 64 times', and 'active 1 year ago'. It is also 'FEATURED ON META'. A yellow banner for the Spanish version of Stack Overflow is visible on the left side of the question area.

StackExchange 2 4 help Search Q&A

stackoverflow Questions Jobs Documentation Beta Tags Users Badges Ask Question

Announcing Stack Overflow Documentation

We started with Q&A. Technical documentation is next, and we need your help. Whether you're a beginner or an experienced developer, you *can* contribute.

I want to help →

Dismiss

How to subset a data frame, apply a function with a for-loop and obtain the result with the name of each subset on it?

asked 1 year ago
viewed 64 times
active 1 year ago

FEATURED ON META

Más de 410 millones de personas hablan español en el mundo. Ven a compartir tu conocimiento.

1 I have a data set where I want to subset it by some conditions as TRT and then evaluate some functions on each one, but instead of repeat the function for every subset I rather to do a for loop. The thing is that I want to get the result but with the name of the subset on it, and I cannot do it

- <http://stackoverflow.com>

Help:

- Recomendaciones de foros:

<http://www.r-project.org/posting-guide.html>

- No hacer preguntas repetidas
- A buenas preguntas, buenas respuestas



[\[Home\]](#)

Download

[CRAN](#)

R Project

[About R](#)

[Contributors](#)

[What's New?](#)

[Mailing Lists](#)

[Bug Tracking](#)

[Conferences](#)

[Search](#)

Posting Guide: How to ask good questions that prompt useful answers

This guide is intended to help you get the most out of the R mailing lists, and to avoid embarrassment. Like many responses posted on the list, it is written in a concise manner. This is not intended to be unfriendly - it is more a consequence of allocating the limited available time and space to technical issues rather than to social niceties.

The list: Remember that R is free software, constructed and maintained by volunteers. They have various reasons for contributing software and participating on the mailing lists, but often have limited time.

Good manners: Remember that customs differ. Some people are very direct. Others surround everything they say with hedges and apologies. Be tolerant. Rudeness is never warranted, but sometimes 'read the manual' *is* the appropriate response. Don't waste time discussing such matters on the list. Ad hominem comments are absolutely out of place.

Questions about statistics: The R mailing lists are primarily intended for questions and discussion

Learning resources:

The screenshot shows the Coursera interface for the 'R Programming' course. At the top, the Coursera logo is on the left, and navigation links for 'Catalog', 'Search catalog', 'Institutions', and 'RM' are on the right. A dark blue banner below the header promotes the 'Data Science Specialization' with an 'Upgrade' button. The left sidebar contains the Johns Hopkins University logo and a list of course navigation links: 'Course Home', 'Course Content', 'Assignments', 'Discussion Forums', and 'Course Info' (which is highlighted with a blue bar). The main content area displays the course title 'R Programming' and its affiliation with Johns Hopkins University. It also mentions that the course is part of a 10-course series, the 'Data Science Specialization'. Below this, there are two columns of information. The left column, titled 'About this Course', provides a detailed description of the course content, including topics like installing software, programming concepts, and practical issues in statistical computing. The right column shows the current session dates (June 27 - July 31) and the upcoming session dates (August 1 - September 5), with buttons for 'Switch sessions' and 'Help Center'.

coursera Catalog Search catalog Institutions RM

Like this course? Become an expert by joining the [Data Science Specialization](#). Upgrade

JOHNS HOPKINS UNIVERSITY

Course Home
Course Content
Assignments
Discussion Forums
Course Info

R Programming

Johns Hopkins University

Part of a 10-course series, the [Data Science Specialization](#)

About this Course

In this course you will learn how to program in R and how to use R for effective data analysis. You will learn how to install and configure software necessary for a statistical programming environment and describe generic programming language concepts as they are implemented in a high-level statistical language. The course covers practical issues in statistical computing which includes programming in R, reading data into R, accessing R packages, writing R functions, debugging, profiling R code, and organizing and commenting R code. Topics in statistical data analysis will provide working

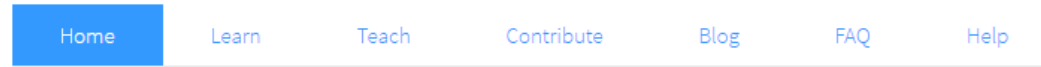
You're currently enrolled in this session:
June 27 – July 31

Upcoming session:
August 1 - September 5

Switch sessions Help Center

- <https://www.coursera.org>

Learning resources:



Learn R, in R.

swirl teaches you R programming and data science
interactively, at your own pace, and right in the R console!

Follow @swirlstats

Got questions? Join our [discussion group](#)!

- <http://swirlstats.com/>

References:

- Hands on programming with R
- Teetor, P. (2011). *R cookbook*. " O'Reilly Media, Inc."
- Adler, J. (2010). *R in a nutshell: A desktop quick reference*. " O'Reilly Media, Inc."
- R programming
- The R book
- R for Data Science