

# IntRoduzione

Andrea Brancaccio, Ph.D

Università di Padova

Anno Accademico 2023/2024

## Table of contents

# Un po' di stoRia

---

- I primi sviluppatori di R sono stati Robert Gentleman e Ross Ihaka del Dipartimento di Statistica dell'Università di Auckland
- R si sviluppa dal linguaggio S
- Fun Fact: il nome deriva dal susseguirsi alfabetico delle lettere
- A partire da 1997 si è costituito un gruppo di studiosi provenienti da ogni parte del mondo denominato “R Development Core Team”
- L' “R Development Core Team” si occupa di sviluppare e aggiornare costantemente R e di diffondere R attraverso meeting e conferenze

# PeRché?

---

## Perché usare R

- R è gratuito, open source e GNU
- RStudio (la di R) no, ma la versione free è più che sufficiente
- Permette la replicabilità dei risultati
- *R community is the best feature of R*

## Software statistico?

“Many users think of R as a statistics system. We prefer to think of it as an environment within which statistical techniques are implemented.”

Cit. <https://www.r-project.org/about.html>

# Come è fatto

Tot - master - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

```
1 # DATA PREPARATION TOL
2 # OTTAVIA, MAGGIO 2023
3 #
4 rm(list = ls())
5 library(lubridate)
6 library(dplyr)
7
8 setwd("H:/./shortcut-targets-by-id/1740KGGGXve6z7MTnN0OcT3W23DXIim")
9
10 name.data.43 = paste0("AdapTo1_43/", # do not change
11                       "to143_2023_05_08.csv") # change according to
12 name.data.52 = paste0("AdapTo1_52/", # do not change
13                       "to152_2023_05_08.csv") # change according to
14 name.data.45 = paste0("AdapTo1_45/", # do not change
15                       "to145_2023_05_08.csv") # change according to
16 name.env = ls()
17 name.env = name.env[grep1("name.data", name.env)]
18
19 for (i in 1:length(name.env)) {
20   assign(gsub("name.", "", name.env[i]),
21         read.csv(get(name.env[i]), header = T, sep = ","))
22 }
23
```

1984 total\_time R Script

Console Terminal Background Jobs

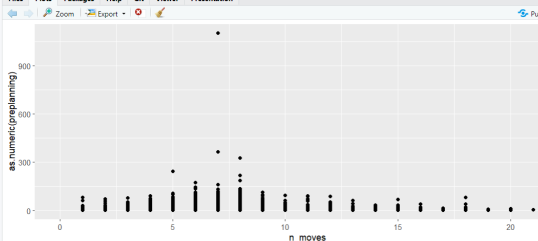
```
R 4.2.3 - H:/./shortcut-targets-by-id/1740KGGGXve6z7MTnN0OcT3W23DXIim/PyCassist/
> ggplot(all.data, aes(x = n_moves, y = as.numeric(preplanning))) + geom_
point()
warning message:
Removed 8 rows containing missing values (geom_point).
>
```

Environment History Connections Tutorial

R - Global Environment

total_time43	378 obs. of 29 variables
total_time45	166 obs. of 44 variables
total_time52	412 obs. of 36 variables
wide.acc	412 obs. of 28 variables
Values	
alleged.trials	num [1:3] 20 35 27
i	12L
name_desc	chr [1:3] "sbj_char43" "sbj_char45" "sbj_char52"
name.data.43	"AdapTo1_43/to143_2023_05_08.csv"
name.data.45	"AdapTo1_45/to145_2023_05_08.csv"
name.data.52	"AdapTo1_52/to152_2023_05_08.csv"
name.env	chr [1:3] "data.43" "data.45" "data.52"
names.print	chr [1:12] "accuracy43" "accuracy45" "accuracy52" "execution43" "executio

Files Plots Packages Help Git Viewer Presentation



# Come è fatto

Tot - master - RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

```
1 # DATA PREPARATION TOL
2 # OTTAVIA, MAGGIO 2023
3 #
4 rm(list = ls())
5 library(lubridate)
6 library(dplyr)
7
8 setwd("H:/./shortcut-targets-by-id/1740KGGGXve6z7MTnN0OcT3W23DXIim")
9
10 name.data.43 = paste0("AdapTo1_43/", # do not change
11                       "to143_2023_05_08.csv") # change according to
12 name.data.52 = paste0("AdapTo1_52/", # do not change
13                       "to152_2023_05_08.csv") # change according to
14 name.data.45 = paste0("AdapTo1_45/", # do not change
15                       "to145_2023_05_08.csv") # change according to
16 name.env = ls()
17 name.env = name.env[grep1("name.data", name.env)]
18
19 for (i in 1:length(name.env)) {
20   assign(gsub("name.", "", name.env[i]),
21         read.csv(get(name.env[i]), header = T, sep = ","))
22 }
23
```

1984 total\_time R Script

Console Terminal Background Jobs

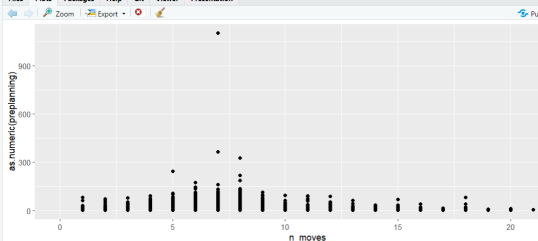
```
R 4.2.3 - H:/./shortcut-targets-by-id/1740KGGGXve6z7MTnN0OcT3W23DXIim/PsyAssist/
> ggplot(all.data, aes(x = n_moves, y = as.numeric(preplanning))) + geom_
point()
warning message:
Removed 8 rows containing missing values (geom_point).
>
```

Environment History Connections Tutorial

R - Global Environment

total_time43	378 obs. of 29 variables
total_time45	166 obs. of 44 variables
total_time52	412 obs. of 36 variables
wide.acc	412 obs. of 28 variables
Values	
alleged.trials	num [1:3] 20 35 27
i	12L
name_desc	chr [1:3] "sbj_char43" "sbj_char45" "sbj_char52"
name.data.43	"AdapTo1_43/to143_2023_05_08.csv"
name.data.45	"AdapTo1_45/to145_2023_05_08.csv"
name.data.52	"AdapTo1_52/to152_2023_05_08.csv"
name.env	chr [1:3] "data.43" "data.45" "data.52"
names.print	chr [1:12] "accuracy43" "accuracy45" "accuracy52" "execution43" "executio

Files Plots Packages Help Git Viewer Presentation



# Con cosa lavoriamo

La gestione delle working directory è una delle cose meno intuitive di R. Rischiare di perdere dei pezzi e di non avere tutto ordinato nelle cartelle. Per risolvere questo problema → si possono usare gli R project (progetti R). Un progetto R crea una sua directory, tutti i file che vengono salvati al suo interno sono sempre accessibili senza bisogno di settare *a mano* tutte le volte la directory.

Sono molto comodi perché:

- 1 Permettono di avere più istanze R aperte contemporaneamente → è possibile lavorare su più progetti contemporaneamente
- 2 Tenendo tutti i file ordinati vi permettono di poter risalire a cosa avete fatto mesi prima (utile quando vi arrivano le revisioni di un paper)

## Creare un progetto R

File → New project:

New Project Wizard

Type 'contributors()' for more  
'citation()' on how to cite R

emos,

# Le basi

## Simboli, variabili, funzioni

Simboli    variabili    funzioni

### Simboli

Operatori:

$3 + 2$     # più  
 $3 - 2$     # meno  
 $3 * 2$     # per  
 $3 / 2$     # diviso  
 $5 > 2$     # maggiore di  
 $2 < 5$     # minore di  
 $5 == 3 + 2$  # uguale  
 $5 != 3$     # diverso

Parentesi e altri simboli

`() [] {} " " : ; ,`

Operatori e parentesi si possono combinare insieme per risolvere equazioni:



# L'ambiente e l'ordine

## Take out the trash

L'ambiente di R dovrebbe essere sempre ordinato

Gli oggetti che non servono più andrebbero eliminati

MA si può anche togliere tutto insieme

```
ls() # lista gli oggetti che sono nell'environment
```

```
rm(A) # rimuove l'oggetto A dall'environment
```

```
rm(list=ls()) # rimuove tutto dall'environment
```

## Save the environment

A volte conviene salvare non solo lo script di R ma anche i risultati che avete ottenuto con quello script

```
save.image("my-computations.RData")
```

Questo ambiente può essere ricaricato nell'ambiente di R:

```
load("my-computations.RData")
```

# Working directories

---

Se decidete di non usare i progetti di R (non è una buona idea, choose your own adventure) dovete conoscere le vostre directory:

`getwd()` # restituisce la working directory in cui siete

`dir()` # restituisce tutti gli oggetti che sono nella wd

Per cambiare la working directory:

`setwd("C:/Users/huawei/OneDrive/Documenti/GitHub/coRso")`