



Universidad de Chile
Facultad de Ciencias Sociales
Carrera de Sociología
Estadística Descriptiva 2023

Laboratorio de Análisis de Datos 1

Sesión 1: Introducción al lenguaje 

Sesión 2: Procesamiento, limpieza y manipulación de datos en  (I)

13 de abril de 2023
Profesor: Rodrigo Asún Inostroza
Prof. Colab.: Andreas Laffert Tamayo
Daniela Olivares Collío



Contenidos LAB 1

Sesión 1: Introducción al lenguaje

- 1.1. Software, Ciencia Abierta y R
- 1.2. R environment: interfaz de RStudio, script, workspace
- 1.3. Ejemplos del uso de R
- 1.4. Prácticas y herramientas de consulta

Sesión 2: Procesamiento, limpieza y manipulación de datos en (I)

- 2.1. Lógica de R Project y flujo de trabajo
- 2.2. R como calculadora
- 2.3. Lógica orientada a objetos (crear objetos, vectores y dataframes)
- 2.4. Tipos de variables de R

1.1. Software, Ciencia Abierta y

1.1. Software, Ciencia Abierta y

¿Qué debo considerar para decidir aprender un nuevo software?

- ¿Qué tipo de datos quiero analizar?
- ¿Cuáles son los costos netos para aprender un lenguaje?
- ¿Cuáles son las herramientas más usadas en mi disciplina?
- ¿Cuáles son las otras herramientas disponibles y cómo se relacionan entre sí?
- Desarrollo del software en el tiempo.





1.1. Software, Ciencia Abierta y

Relevancia del uso de software en investigación social

- El análisis de datos sociales hoy: tendencias emergentes en los procesos de producción de información social.
- Es central manejar herramientas altamente flexibles.
- Diferentes alternativas de softwares para análisis de datos.
- El software como decisión metodológica: conocer y controlar potencialidades y límites de nuestra herramienta



1.1. Software, Ciencia Abierta y

Open Science



Fuente: UNESCO

- La ciencia abierta busca hacer que la investigación científica sea más **accesible, transparente, colaborativa y eficiente**, lo que puede tener un impacto positivo en la calidad y relevancia de la investigación, así como en la sociedad en general.
- Busca fomentar la **colaboración**, el **intercambio** de información y la **difusión** de conocimientos.
- Implica la publicación de datos, artículos, informes y otros tipos de resultados de investigación en formatos que sean de acceso libre y que permitan la reutilización de los mismos.

1.1. Software, Ciencia Abierta y

Open Science



NOTICIAS | CULTURA+CIUDAD

CULTURA | CIENCIA



Profesor de universidad chilena es reconocido por aporte a misión de la NASA

por El Mostrador Cultura | 20 mayo, 2021



Chihau Chau es miembro del Departamento de Informática de la Universidad Técnica Federico Santa María y mentor del Instituto 3IE. Un código escrito por él fue usado en una reciente expedición a Marte.



Chihau Chau es miembro del Departamento de Informática de la Universidad Técnica Federico Santa María.

Achievements

 Ruby 



 You have a new badge!

You contributed code to repositories used in the Mars 2020 Helicopter Mission and now have a badge for it. Thank you!

OK, got it

Opt out in settings.

tributio

r May



arn how

Fuente: [UTFSM](#), [El Mostrador](#), [Biobio](#)



1.1. Software, Ciencia Abierta y

¿Por qué optar por software libre?

Definición: “El software libre es una cuestión de libertad, no de precio. (...) Con software libre nos referimos a la libertad que tienen los usuarios para ejecutar, copiar, distribuir, estudiar, modificar y mejorar el software.” (Stallman, 2004).

- Un gran punto a favor es no tener que pagar licencia.
- No es sólo cuestión de “costo”: ejemplos de uso limitante de software pagado: Técnica como caja negra.
- Herramientas computacionales e investigación científica: deberíamos poder tener control sobre todos los componentes de nuestro proceso investigativo.

1.1. Software, Ciencia Abierta y

Open Science y

- Software libre (freeware): Licencia Pública General (General Public License).
- Distribución gratuita de su versión básica y funciones especializadas.
- R Development Core Team y CRAN (Comprehensive R Archive Network) como espacio de desarrollo y distribución del software.
- Actualización pública, gratuita y permanente de versión básica y especializada.
- Contribución de comunidades científicas activas alrededor del mundo desarrollan herramientas de procesamiento, visualización y análisis de datos.

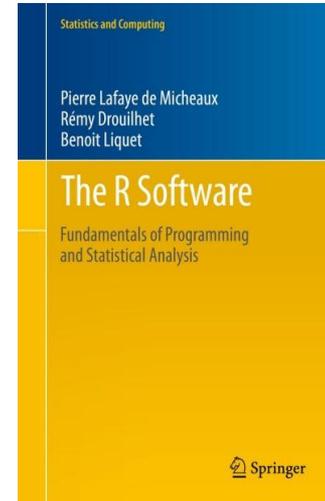


Fuente: UNESCO

1.1. Software, Ciencia Abierta y R

¿Por qué elegir ?

- Gratis
- Open Source
- Multiplataforma (Windows, MacOS, Linux)
- De evolución rápida (mediante paquetes)
- Potente en la manipulación de datos, cálculo y gráficos.
Procedimientos eficientes para el tratamiento de datos y su almacenamiento.
- Sistema de documentación integrado y muy bien concebido.
- RStudio: interfaz gráfica de R.



1.2. environment: interfaz de Studio[®], script, workspace

1.2. R environment: interfaz de script, workspace



The screenshot shows the R Studio interface with several key components highlighted by callouts:

- Entorno de Trabajo**: Points to the Environment pane on the right, which shows the Global Environment.
- Editor de sintaxis**: Points to the source editor on the left, which contains R code for setting the working directory and installing/loading packages.
- Visualizador de resultados**: Points to the Console pane at the bottom, which shows the output of the commands executed in the script.
- Archivos, gráficos, paquetes, ayuda**: Points to the Packages pane on the right, which lists installed and available packages.

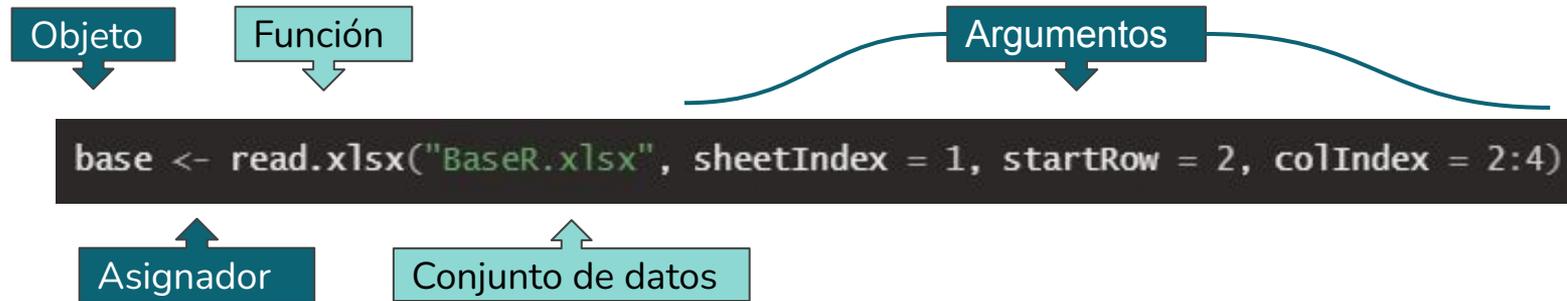
```
1 #----1. Fijar Directorio de Trabajo, instalar y cargar paquetes----#
2
3 setwd("/media/chuikov/Temporal")
4
5 install.packages("ggplot2")
6 install.packages("dygraphs")
7 install.packages(c("haven", "tidyverse", "gapminder", "gganimate", "gifski"))
8
9 library(haven)
10 library(tidyverse)
11 library(ggplot2)
12 library(gapminder)
13 library(gganimate)
14 library(dygraphs)
15 library(gifski)
16
17
18 a<-10
19
```

Console Terminal x

```
/media/chuikov/Temporal/ >
> library(tidyverse)
> library(ggplot2)
> library(gapminder)
> library(gganimate)
> library(dygraphs)
> library(gifski)
>
```

Name	Description	Version
<input type="checkbox"/> askpass	Safe Password Entry for R, Git, and SSH	1.1
<input type="checkbox"/> backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.3
<input type="checkbox"/> base64enc	Tools for base64 encoding	0.1-3
<input type="checkbox"/> bazar		0.11
<input type="checkbox"/> bitops		0-6
<input type="checkbox"/> broom		1.1
<input type="checkbox"/> callr		2.0
<input type="checkbox"/> caTools	Tools: moving window statistics, GIF, Base64, ROC AUC, etc.	1.17.1.2
<input type="checkbox"/> cellranger	Translate Spreadsheet Cell Ranges to Rows and Columns	1.1.0
<input type="checkbox"/> checkmate	Fast and Versatile Argument	1.9.1

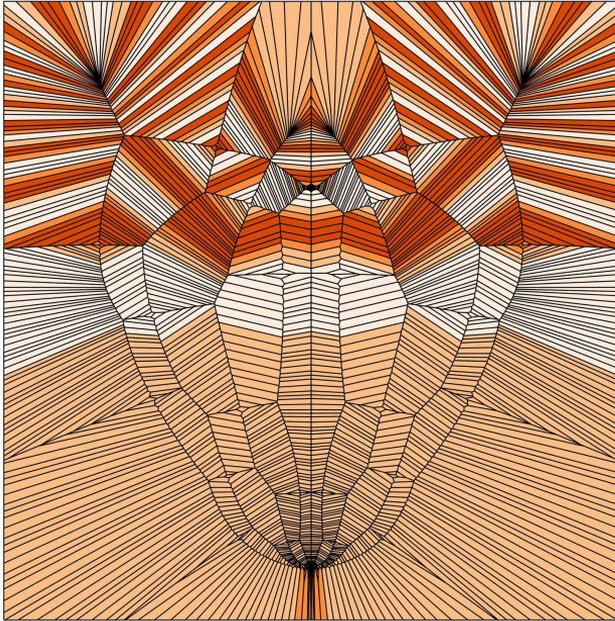
1.2. R environment: interfaz de Studio[®] script, workspace



- A un objeto se le asigna el resultado de una función que se ejecuta sobre un conjunto de datos especificado y una serie de configuraciones particulares.
- Esta es la estructura de un comando básico del lenguaje R.
- Hay paquetes que tienen ciertas especificidades y hay algunos “atajos” a esta estructura.

1.3. Ejemplos del uso de

1.3. Ejemplos del uso de

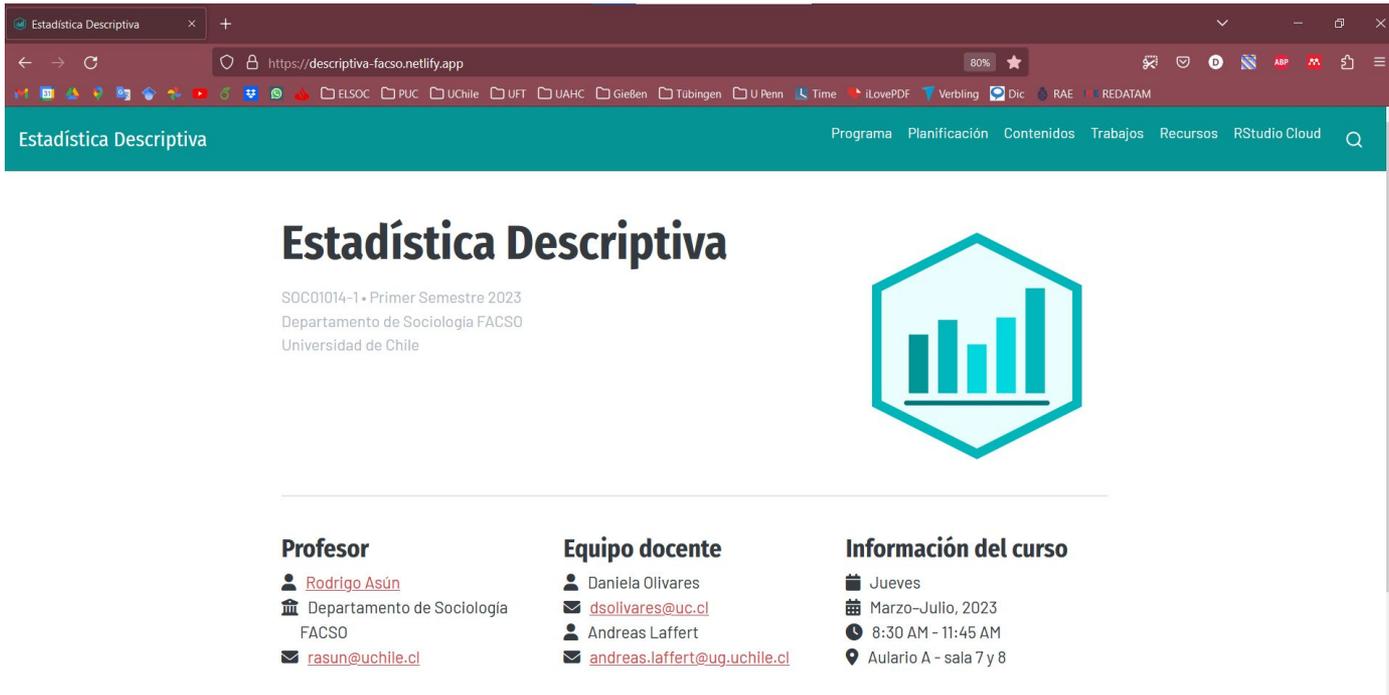


```
library(ggforce)
library(dplyr)

tibble(t=seq(-6,6,8e3)) %>%
  mutate(x=sin(5*t)^2*2^cos(cos(8*t)),y=sin(sin(5*t))*cos(7*t)^2) %>%
  ggplot(aes(x,y,fill=sin(x^6))) +
  geom_voronoi_tile(color=1) +
  scale_fill_fermenter(guide="none",palette=7) +
  coord_flip() +
  theme_void()
```

Fuente: [Georgios Karamanis en Twitter](#)

1.3. Ejemplos del uso de



Estadística Descriptiva

Programa Planificación Contenidos Trabajos Recursos RStudio Cloud

Estadística Descriptiva

SOC01014-1 • Primer Semestre 2023
Departamento de Sociología FACSO
Universidad de Chile



Profesor  Rodrigo Asún  Departamento de Sociología FACSO  rasun@uchile.cl	Equipo docente  Daniela Olivares  dsolivares@uc.cl  Andreas Laffert  andreas.laffert@ug.uchile.cl	Información del curso  Jueves  Marzo-Julio, 2023  8:30 AM - 11:45 AM  Aulario A - sala 7 y 8
--	--	---

1.4. Prácticas y herramientas de consulta

1.4. Prácticas y herramientas de consulta

- CRAN
- Stack overflow
- R Community
- Chat GPT
- Otros



Jesse Maegan
@kierisi

Following

My #rstats learning path:

1. Install R
2. Install RStudio
3. Google "How do I [THING I WANT TO DO] in R?"

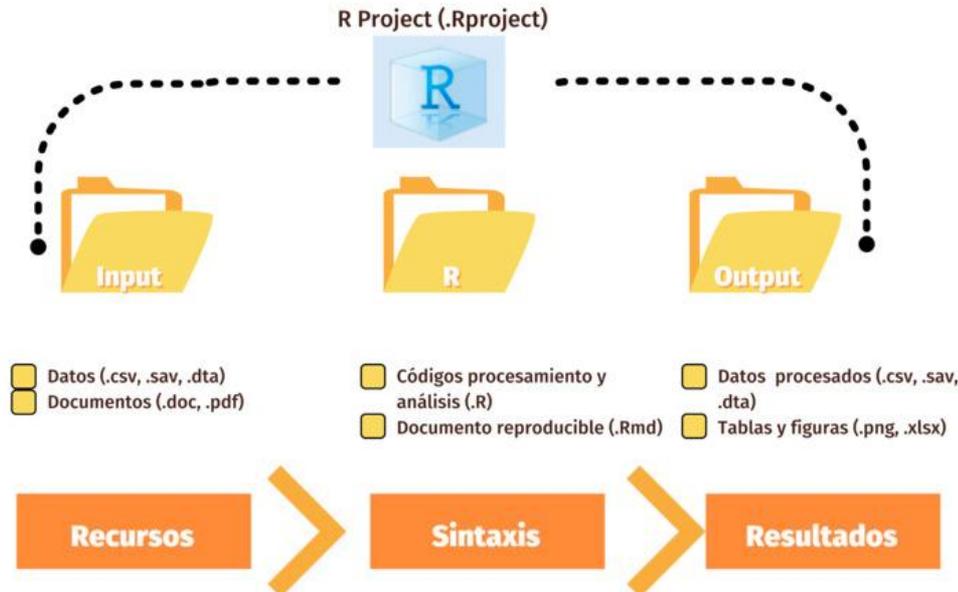
Repeat step 3 ad infinitum.

7:19 AM - 18 Aug 2017

2.1. Lógica de R Project y flujo de trabajo

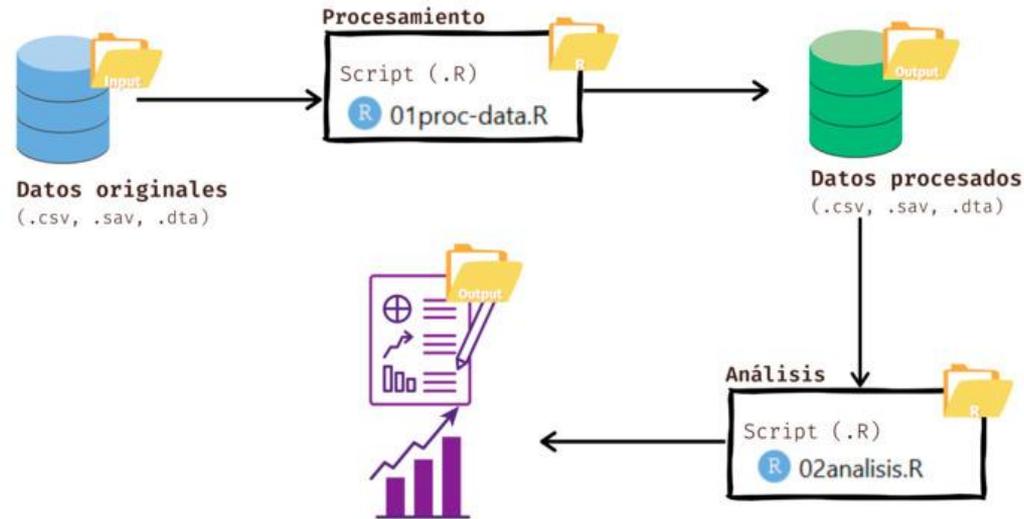
2.1. Lógica de R Project y flujo de trabajo

R Project



2.1. Lógica de R Project y flujo de trabajo

Flujo de trabajo





Universidad de Chile
Facultad de Ciencias Sociales
Carrera de Sociología
Estadística Descriptiva 2023

Laboratorio de Análisis de Datos 1

Sesión 1: Introducción al lenguaje 

Sesión 2: Procesamiento, limpieza y manipulación de datos en  (I)

13 de abril de 2023
Profesor: Rodrigo Asún Inostroza
Prof. Colab.: Andreas Laffert Tamayo
Daniela Olivares Collío