

# Taller 1: Exploración inicial con R y Rstudio

José Daniel Conejeros

Magíster(c) Sociología UC - Licenciado en Ciencias Sociales

Junio 2020

🔗 Código del taller 📩 jdconejeros@uc.cl 🌐 JDConejeros

linkedin joseconejerosp R<sup>g</sup> Jose\_Conejeros

# ¿Quiénes somos?



**Guía del taller:** Estoy en proceso de titulación del Magíster de Sociología UC y actualmente participo en proyectos de investigación sobre estratificación y justicia social. Además colaboro en docencia en los cursos de análisis en el Magíster de Políticas Públicas de la Escuela de Gobierno UC. Me gusta pasear en bicicleta, ir al cine, leer libros lateros y el jugo de mango. Correo: [jdconejeros@uc.cl](mailto:jdconejeros@uc.cl)



**Apoyo técnico:** Soy titulada en Ciencia Política UC y Certificada en Psicología, actualmente estoy a cargo de la parte administrativa del taller y preparándome para la prueba de ingreso de la Academia Diplomática. Me gusta mucho hacer deporte, bailar, leer, escribir y soy fanática de los helados. Correo: [cplemus@uc.cl](mailto:cplemus@uc.cl)

¿Y ustedes?

# Recomendaciones para una buena convivencia



- Trato respetuoso
- Puntualidad
- Uso del micrófono
- Trabajo colaborativo
- Comunicación

"Esperamos que estas recomendaciones sean consideradas de forma positiva  
y orientadas al bien común."



# ¿Por qué R y Rstudio?

Apunta a distintos **perfles** (usuarios, desarrolladores), permite el uso de **paquetes**, cuenta con herramientas de **visualización**, tiene un **flujo** completo de trabajo, existe una **comunidad muy activa** y permite el desarrollo de **grupos de interés**.





# Instalar R y RStudio

Primero tienes que descargar R desde <https://cran.r-project.org/>. Recuerda seleccionar la versión apropiada a tu computador:

**The Comprehensive R Archive Network**

**Download and Install R**

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 (2020-06-06, See Things Now) [R-4.0.1.targz](#), 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](#)

**Questions About R**

- 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.



# El instalador para MAC

## R for Mac OS X

This directory contains binaries for a base distribution and packages to run on Mac OS X (release 10.6 and above). Mac OS 8.6 to 9.2 (and Mac OS X 10.1) are no longer supported but you can find the last supported release of R for these systems (which is R 1.7.1) [here](#). Releases for old Mac OS X systems (through Mac OS X 10.5) and PowerPC Macs can be found in the [old](#) directory.

Note: CRAN does not have Mac OS X systems and cannot check these binaries for viruses. Although we take precautions when assembling binaries, please use the normal precautions with downloaded executables.

Package binaries for R versions older than 3.2.0 are only available from the [CRAN archive](#) so users of such versions should adjust the CRAN mirror setting (<https://cran-archive.r-project.org>) accordingly.

R 4.0.1 "See Things Now" released on 2020/06/06

Please check the MD5 checksum of the downloaded image to ensure that it has not been tampered with or corrupted during the mirroring process. For example type

md5 R-4.0.1.pkg

in the *Terminal* application to print the MD5 checksum for the R-4.0.1.pkg image. On Mac OS X 10.7 and later you can also validate the signature using

pkgutil --check-signature R-4.0.1.pkg

**R-4.0.1.pkg** (notarized and signed)  
SHA1-hash: c21adbc6e3362cac24969dc9916c3bfffde062b8  
(ca. 84MB)

### Latest release:

**R 4.0.1** binary for macOS 10.13 (High Sierra) and higher, signed and notarized package. Contains R 4.0.0 framework, R.app GUI 1.72 in 64-bit for Intel Macs, Tcl/Tk 8.6.6 X11 libraries and Texinfo 6.7. The latter two components are optional and can be omitted when choosing "custom install", they are only needed if you want to use the `tcltk` R package or build package documentation from sources.

Note: the use of X11 (including `tcltk`) requires [XQuartz](#) to be installed since it is no longer part of OS X. Always re-install XQuartz when upgrading your macOS to a new major version.

**Important:** this release uses Xcode 10.1 and GNU Fortran 8.2. If you wish to compile R packages from sources, you will need to download and GNU Fortran 8.2 - see the [tools](#) directory.

[NEWS](#) (for Mac GUI)

News features and changes in the R.app Mac GUI

[Mac-GUI-1.70.tar.gz](#)

MD5-hash: b1ef5f285524640680a22965bb880008

Sources for the R.app GUI 1.70 for Mac OS X. This file is only needed if you want to join the development of the GUI, it is not intended for regular users. Read the INSTALL file for further instructions.

Note: Previous R versions for El Capitan can be found in the [el-capitan/base](#) directory.

Binaries for legacy OS X systems.

## "Debes descargar y seguir los pasos de instalación"

Click aquí para instalar

# El instalador para WINDOWS



## R for Windows

Subdirectories:

[base](#)

Binaries for base distribution. This is what you want to [install R for the first time](#).

[contrib](#)

Binaries of contributed CRAN packages (for R >= 2.13.x; managed by Uwe Ligges). There is also information on [third party software](#) available for CRAN Windows services and corresponding environment and make variables.

[old\\_contrib](#)

Binaries of contributed CRAN packages for outdated versions of R (for R < 2.13.x; managed by Uwe Ligges).

[Rtools](#)

Tools to build R and R packages. This is what you want to build your own packages on Windows, or to build R itself.

Please do not submit binaries to CRAN. Package developers might want to contact Uwe Ligges directly in case of questions / suggestions related to Windows binaries.

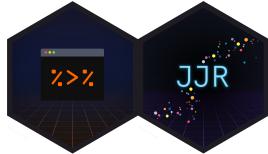
You may also want to read the [R FAQ](#) and [R for Windows FAQ](#).

Note: CRAN does some checks on these binaries for viruses, but cannot give guarantees. Use the normal precautions with downloaded executables.

"Debes descargar y seguir los pasos de instalación"

Click aquí para instalar

# Descargar e instalar Rstudio



RStudio Desktop 1.3.959 - [Release Notes](#)

1. Install R. Rstudio requires [R 3.0.1+](#).
2. Download RStudio Desktop. Recommended for your system:



Requires macOS 10.13+ (64-bit)



## All Installers

Linux users may need to [import RStudio's public code-signing key](#) prior to installation, depending on the operating system's security policy.

RStudio requires a 64-bit operating system. If you are on a 32 bit system, you can use an [older version of RStudio](#).

[Links de Descarga](#)

OS	Download	Size	SHA-256
Windows 10/8/7	<a href="#"> RStudio-1.3.959.exe</a>	171.41 MB	3d493ae5
macOS 10.13+	<a href="#"> RStudio-1.3.959.dmg</a>	148.57 MB	7c5b695d

[Click aquí para instalar](#)

# Trabajar en la nube Rstudio Cloud



Si no queremos/podemos instalar R en nuestros equipos, tenemos la opción de usar su versión en la nube. Para empezar a trabajar con RStudioCloud necesitamos **crear una cuenta**.

The screenshot displays the RStudio Cloud interface. On the left, a sidebar menu includes options like 'Your Workspace / Mi Proyecto', 'File', 'Edit', 'Code', etc., and links for 'Spaces', 'New Space', 'Learn', 'Guide', 'What's New', 'Primers', 'Cheat Sheets', 'Feedback and Questions', 'Info', 'Plans & Pricing', 'Terms and Conditions', and 'System Status'. The main workspace area has tabs for 'Console', 'Terminal', and 'Jobs'. The 'Console' tab shows the R startup message:

```
R version 4.0.0 (2020-04-24) -- "Arbor Day"
Copyright (C) 2020 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.

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.
```

The 'Environment' tab in the top right shows an empty global environment. The 'Files' tab at the bottom shows a 'project' folder containing '.Rhistory' (0 B, Jun 11, 2020, 10:58 PM) and 'project.Rproj' (205 B, Jun 11, 2020, 10:58 PM).

# Esto es R



Es un lenguaje y entorno para computación estadística y gráficos. Está disponible como software libre.

R version 3.6.0 (2019-04-26) -- "Planting of a Tree"  
Copyright (C) 2019 The R Foundation for Statistical Computing  
Platform: x86\_64-w64-mingw32/x64 (64-bit)

R es un software libre y viene sin GARANTIA ALGUNA.  
Usted puede redistribuirlo bajo ciertas circunstancias.  
Escriba 'license()' o 'licence()' para detalles de distribución.

R es un proyecto colaborativo con muchos contribuyentes.  
Escriba 'contributors()' para obtener más información y  
'citation()' para saber cómo citar R o paquetes de R en publicaciones.

Escriba 'demo()' para demostraciones, 'help()' para el sistema on-line de ayuda,  
o 'help.start()' para abrir el sistema de ayuda HTML con su navegador.  
Escriba 'q()' para salir de R.

> |

R Gui (64-bit) R Console

R version 3.6.0 (2019-04-26) -- "Planting of a Tree"  
Copyright (C) 2019 The R Foundation for Statistical Computing  
Platform: x86\_64-apple-darwin15.6.0 (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.

During startup - Warning messages:  
1: Setting LC\_CTYPE failed, using "C"  
2: Setting LC\_COLLATE failed, using "C"  
3: Setting LC\_TIME failed, using "C"  
4: Setting LC\_MESSAGES failed, using "C"  
5: Setting LC\_MONETARY failed, using "C"  
[R.app GUI 1.70 (7657) x86\_64-apple-darwin15.6.0]

WARNING: You're using a non-UTF8 locale, therefore only ASCII characters will work.  
Please read R for Mac OS X FAQ (see Help) section 9 and adjust your system.

# Esto es Rstudio



Es un entorno de desarrollo integrado (IDE) para R. Incluye varios elementos que facilitan las tareas durante el flujo de trabajo.

## Documentos y apps

RStudio desarrolla herramientas gratuitas y abiertas para R. Su entorno de desarrollo integrado (IDE) facilita el análisis de datos con R.

## Escribe tu código

The screenshot shows the RStudio interface with several callout boxes explaining its features:

- Abrir en una nueva ventana**: Open in new window
- Guardar**: Save
- Encontrar y reemplazar**: Find and replace
- Compilar como cuaderno**: Compile as notebook
- Ejecutar el código seleccionado**: Execute selected code
- Esquema del archivo**: File structure
- Historial de comandos**: Command history
- Importar datos con asistente**: Import data with wizard
- Mostrar presentaciones de diapositivas RPres**: Show RPres presentations
- Archivo > Nuevo archivo > Presentación R**: New file > R presentation

**Soporte R**

- También ofrece:** muchos paquetes R (e.g. tidyverse, sparklyr, ggplot2, dplyr), incluidos **Shiny** (crea aplicaciones web sencillas en R) y **R Markdown** (te permite convertir tus análisis en documentos, informes, presentaciones y paneles de alta calidad; compartirlos y reproducirlos).
- Shiny** logo
- markdown** logo

**Escribe tu código**

- Cursor del usuario**: User cursor
- Ejecutar el código previo**: Execute previous code
- Ejecutar el código completo**: Execute full code
- Esquema del archivo**: File structure
- Selección: Alt + arrastre del mouse.**: Selection: Alt + drag mouse
- Resulado de sintaxis.**: Syntax results
- Autocompletar mediante tabulación:** Autocomplete via tabulation
- nombres de funciones, rutas de archivos, argumentos y más.**: Function names, file paths, arguments, etc.
- get\_dgit <-function() {**: Function definition
- ("num" %% (10 ^ n))**: Modulo operation
- %% (10 ^ (n - 1))**: Modulo operation
- foo**: Function name
- for (i in 1:n) {**: For loop
- foo <- function(x) x + 1**: Function assignment
- force**: Force evaluation
- [base]**: Base environment
- Saltar a la función en el archivo**: Jump to function in file
- Cambia el tipo de archivo**: Change file type
- Presiona ↑ para ver el historial de comandos**: Press ↑ to view command history
- Maximiza/minimiza los paneles**: Maximize/minimize panels
- Arrstra los límites**: Drag panel limits

**Soporte R**

- Carga el área de trabajo**: Load workspace
- Guarda el área de trabajo**: Save workspace
- Borra todos los objetos**: Delete all objects
- Buscador del área de trabajo**: Workspace search
- Elegir el entorno para mostrar**: Choose environment to show
- Mostrar objetos como lista o cuadrícula**: Show objects as list or grid
- Data**
  - iris**: 150 obs. of 5 variables
  - Values**: 1
  - Functions**: foo function (x)
- Muestra los objetos guardados por tipo, con una breve descripción**: Show objects by type, with brief description
- Ver en el visor de datos**: View in data viewer
- Ver el código fuente de la función**: View function source code

**Files** tab

- New Folder**, **Upload**, **Delete**, **Rename**, **More**
- Crear una carpeta**: Create a folder
- Subir un archivo**: Upload file
- Eliminar un archivo**: Delete file
- Renombrar un archivo**: Rename file
- Export...**
- Set As Working Directory**
- Co To Working Directory**

**Console** tab

- Presiona ↑ para ver el historial de comandos**: Press ↑ to view command history

**Directorio de trabajo**: Working directory

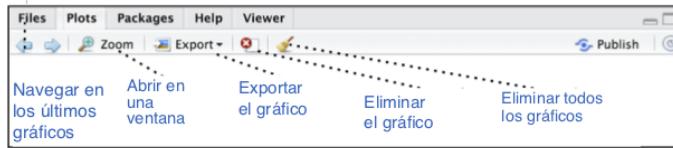
**Maximiza/minimiza los paneles**: Maximize/minimize panels

**Arrstra los límites**: Drag panel limits

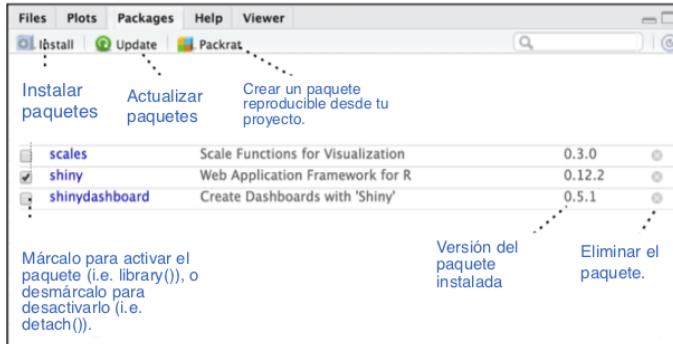
# Configurar nuestro espacio de trabajo



## GRÁFICOS



## PAQUETES



## AYUDA





# ¿Qué son los objetos?

Estructuras de datos que contienen información:

```
objeto ← c("info1", "info2", "info3")
objeto
```

```
## [1] "info1" "info2" "info3"
```

```
ventas_chocolate ← c(254, 203, 182, 50)
ventas_chocolate
```

```
## [1] 254 203 182 50
```

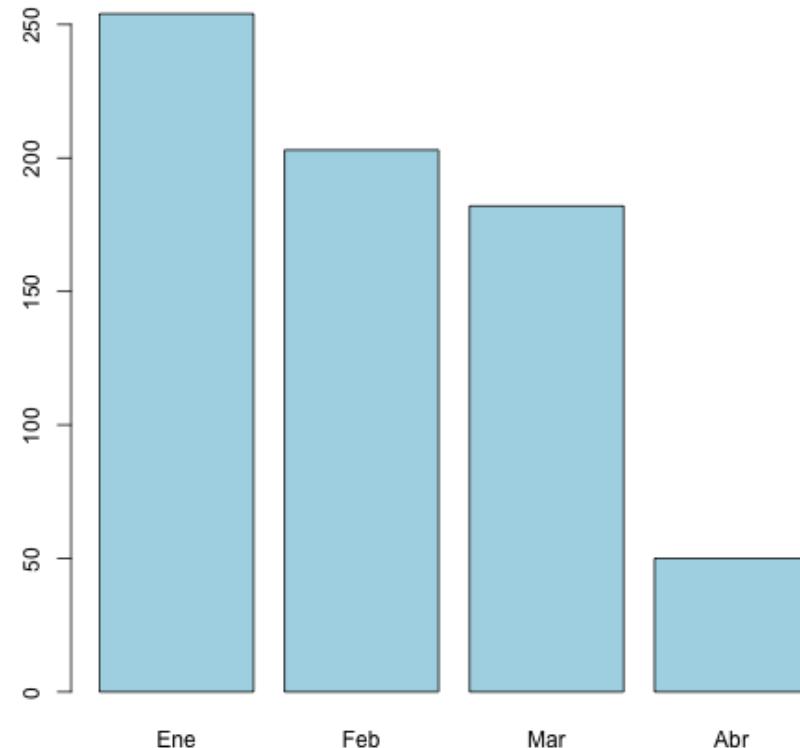
```
meses ← c("Ene", "Feb", "Mar", "Abr")
meses
```

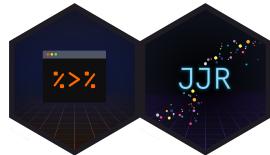
```
## [1] "Ene" "Feb" "Mar" "Abr"
```

# ¿Qué son los objetos?



```
barplot ← barplot(ventas_chocolate, names.arg = meses, col="lightblue")
```





# Vectores, matrices, listas

```
# Esto es un vector
nombre  ← c("José", "Constanza")

edad    ← c(29, 24)

altura  ← c(1.67, 1.65)

matriz  ← cbind(nombre, edad, altura)
matriz

##      nombre      edad  altura
## [1,] "José"     "29"  "1.67"
## [2,] "Constanza" "24"  "1.65"

lista   ← c(matriz, edad, nombre)
lista

##  [1] "José"          "Constanza"      "29"           "24"           "1.67"         "1.65"
##  [7] "29"            "24"            "José"          "Constanza"
```



# Funciones y argumentos

Ejecutan una acción sobre nuestros datos. Algunas requieren **inputs** (argumentos) que van dentro del paréntesis.

```
?sum  
  
sum( ... , na.rm = FALSE)  
  
Arguments  
...           numeric or complex or logical vectors.  
na.rm        logical. Should missing values (including NaN) be removed?
```

```
edad
```

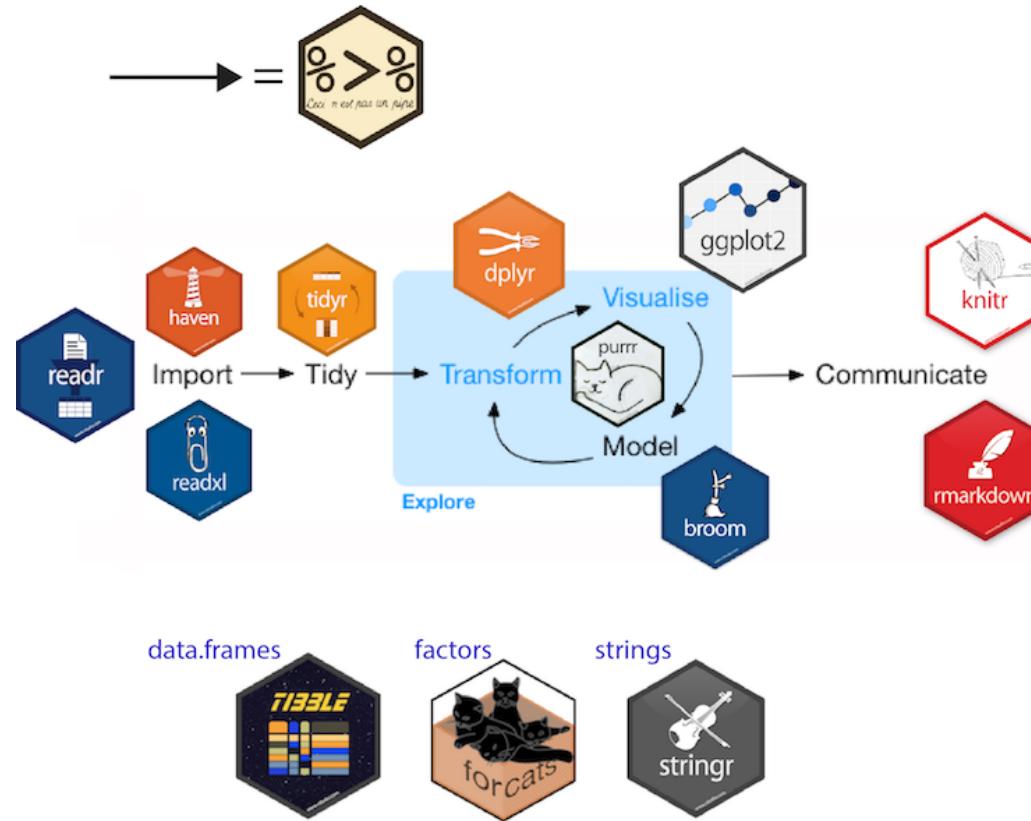
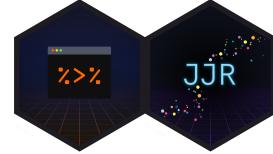
```
## [1] 29 24
```

```
sum(edad)
```

```
## [1] 53
```

Objetos -> Valores -> Funciones

# Uso de librerías o Paquetes



Puedes revisar las principales librerías dando [click aquí](#)

# ¿Cómo instalar librerías o paquetes?



Podemos llamar librerías desde **CRAN** (Comprehensive R Archive Network). Repositorio oficial de paquetes de R (debe ser aprobados).

```
install.packages("nombrepaquete")
```

Podemos llamar librerías desde **GitHub** que es una plataforma de desarrollo de software.

```
install.packages("remotes")
remotes::install_github("usuario/paquete")
```

```
# Calculamos el promedio  
media ← sum(base$casos)/nrow(base)  
  
media  
  
mean(base$casos)  
  
#Lo agregamos como columna a nuestra base de  
base$dif ← (base$casos - media)  
  
head(base, n=16)
```

```
## [1] 9630.75  
  
## [1] 9630.75  
  
##          region  casos      dif  
## 1  Arica y Parinacota  967 -8663.75  
## 2        Tarapacá 3696 -5934.75  
## 3    Antofagasta 3753 -5877.75  
## 4       Atacama  389 -9241.75  
## 5     Coquimbo 1345 -8285.75  
## 6   Valparaíso 5865 -3765.75  
## 7 Metropolitana 124135 114504.25  
## 8      O'Higgins 1552 -8078.75  
## 9        Maule 2560 -7070.75  
## 10       Ñuble 1673 -7957.75  
## 11       Biobío 2974 -6656.75  
## 12     Araucanía 2535 -7095.75  
## 13       Los Ríos  462 -9168.75  
## 14     Los Lagos 1062 -8568.75  
## 15       Aysén   22 -9608.75  
## 16 Magallanes 1102 -8528.75
```

```
# Calculamos las diferencias al cuadrado  
base$dif_2 <- (base$dif^2)  
  
head(base, n=16)
```

	##	region	casos	dif	dif_2
## 1	Arica y Parinacota	967	-8663.75	75060564	
## 2	Tarapacá	3696	-5934.75	35221258	
## 3	Antofagasta	3753	-5877.75	34547945	
## 4	Atacama	389	-9241.75	85409943	
## 5	Coquimbo	1345	-8285.75	68653653	
## 6	Valparaíso	5865	-3765.75	14180873	
## 7	Metropolitana	124135	114504.25	13111223268	
## 8	O'Higgins	1552	-8078.75	65266202	
## 9	Maule	2560	-7070.75	49995506	
## 10	Ñuble	1673	-7957.75	63325785	
## 11	Biobío	2974	-6656.75	44312321	
## 12	Araucanía	2535	-7095.75	50349668	
## 13	Los Ríos	462	-9168.75	84065977	
## 14	Los Lagos	1062	-8568.75	73423477	
## 15	Aysén	22	-9608.75	92328077	
## 16	Magallanes	1102	-8528.75	72739577	

```

# Estimamos
base$resultado ← sum(base$dif_2)/(16-1)

# Estimamos la raíz cuadrada
base$resultado ← sqrt(base$resultado)

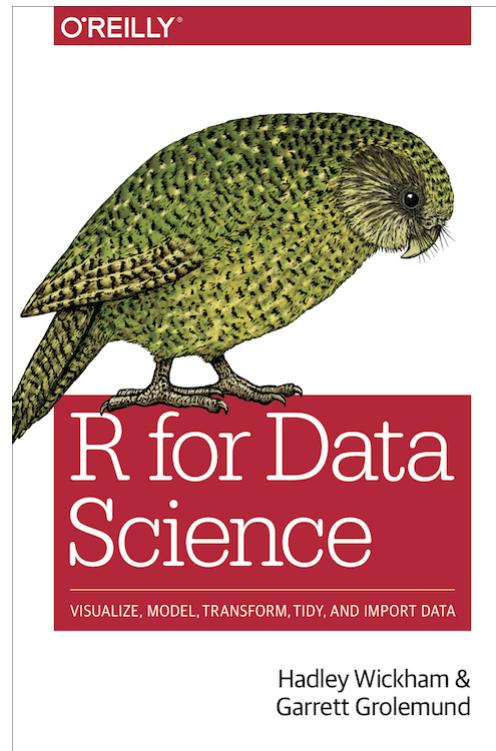
# Vemos nuestro resultado
head(base, n=16)

```

	##	region	casos	dif	dif_2	resultado
## 1	Arica y Parinacota	967	-8663.75	75060564	30572.43	
## 2	Tarapacá	3696	-5934.75	35221258	30572.43	
## 3	Antofagasta	3753	-5877.75	34547945	30572.43	
## 4	Atacama	389	-9241.75	85409943	30572.43	
## 5	Coquimbo	1345	-8285.75	68653653	30572.43	
## 6	Valparaíso	5865	-3765.75	14180873	30572.43	
## 7	Metropolitana	124135	114504.25	13111223268	30572.43	
## 8	O'Higgins	1552	-8078.75	65266202	30572.43	
## 9	Maule	2560	-7070.75	49995506	30572.43	
## 10	Ñuble	1673	-7957.75	63325785	30572.43	
## 11	Biobío	2974	-6656.75	44312321	30572.43	
## 12	Araucanía	2535	-7095.75	50349668	30572.43	
## 13	Los Ríos	462	-9168.75	84065977	30572.43	
## 14	Los Lagos	1062	-8568.75	73423477	30572.43	
## 15	Aysén	22	-9608.75	92328077	30572.43	
## 16	Magallanes	1102	-8528.75	72739577	30572.43	

# Referencias

## R para Ciencia de Datos



Versión en español (aún en desarrollo).

## Otros manuales



## Resolver dudas:



# Sobre la comunidad de R en Chile:

R-Ladies Santiago +  
Chicas Poderosas Chile

INTRODUCCIÓN AL  
PROCESAMIENTO Y VISUALIZACIÓN  
DE DATOS USANDO R



# ¡Gracias!

Presentación generada en:



✉ jdconejeros@uc.cl  joseconejeros

 JDConejeros  Jose\_Conejeros