

Página de la materia :

http://www.dm.uba.ar/materias/analisis_de_datos/2013/2/

PROGRAMA

1. Introducción al R y R-Commander. Creación de un directorio de trabajo separado. Presentación de distintos tipos de gráficos. Objetos en R. Lenguaje: expresiones y asignaciones. Lectura y escritura de estructuras sencillas de datos: funciones scan, read.table, read.csv. Lectura y escritura de funciones.
2. Histogramas. Reglas para cantidad de intervalos de clase: Dixon&Kronmal, Velleman, Sturges. Reglas para la longitud de intervalos: Scott y Friedman&Diaconis. Comparación. Modificación de la cantidad de clases.
3. Diagramas Tallo-Hoja. Esquema básico. Profundidades y cantidad de hojas por tallo. Resistencia.
4. Estadísticos de orden. Cuantiles en distribuciones continuas y discretas. Cuantiles en conjuntos de datos. Funciones sort, sort.list, rank, rev, order, min, max, median, summary, quantile. Diferentes formas de cálculo de los cuantiles. Gráfico cuantil-cuantil.
5. Medidas resumen. Media muestral. Varianza muestral. Desvío. Mediana. Cuartos. Distancia Inter cuartos. Identificación de valores atípicos.
6. Boxplots. Construcción. Resistencia. Interpretación. Puntos de corte para outliers en varias distribuciones, simétricas y asimétricas. Comparación de lotes.
7. Gráfico de dispersión versus nivel. Construcción. Escalera de potencias de Tukey. Justificación del gráfico.
8. Transformaciones de potencia. Definición. Ventajas. Propiedades. Transformaciones apareadas: procedimiento, interpretación.
9. Simetrización de datos. Gráfico de transformación para simetría. Justificación.
10. Regresión lineal simple. Modelo. Interpretación. Estimación. Predicción. Cuadrados mínimos. Supuestos. Residuos. Leverage. Outliers. Puntos influyentes
11. Rectas resistentes. Recta resistente a partir de tres grupos. Mínimos absolutos (L1). Mínima mediana cuadrática (LMS). Otros métodos alternativos.

BIBLIOGRAFÍA:

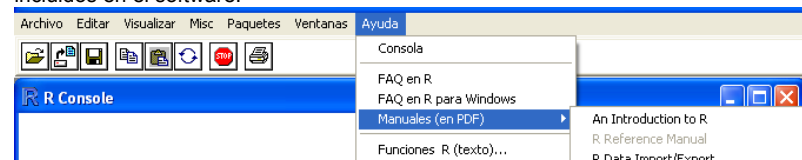
Básico: UNDERSTANDING ROBUST AND EXPLORATORY DATA ANALYSIS. Hoaglin, Mosteller, Tukey (1983). ISBN 0-471-09777-2. Wiley. Nuestro libro de cabecera UREDA.

Complementario: The Elements of Graphing Data. Cleveland, W. S. (1994)
Belmont, CA: Wadsworth.

SOFTWARE

R - R Commander

- **El programa R:** Desde el Help del menú se puede acceder a los manuales incluidos en el software.



Estos están escritos por el grupo de desarrollo de R. Algunos son largos y específicos, pero el manual "An Introduction to R" es una buena fuente de información útil.

- **Documentación:**

Para R

El sitio de CRAN (Comprehensive R Archive Network) <http://cran.r-project.org/>

tiene manuales <http://cran.r-project.org/manuals.html>

y contribuciones de los usuarios <http://cran.r-project.org/other-docs.html>.

Algunas son en castellano, como la traducción de "An Introduction to R". Allí también se encuentran

W. J. Owen 2010. **The R Guide**
<http://cran.r-project.org/doc/contrib/Owen-TheRGuide.pdf>

guía de la que he tomado material;

R reference card by Tom Short es un resumen de 4 páginas muy útil.

Emmanuel Paradis 2003. **R para Principiantes**. En castellano
http://cran.r-project.org/doc/contrib/rdebuts_es.pdf

Para R commander

La página para bajar el software

<http://cran.r-project.org/web/packages/Rcmdr/index.html>

Manuales

<http://cran.r-project.org/doc/contrib/Karp-Rcommander-intro.pdf> ## en inglés

<http://cran.r-project.org/doc/contrib/Saez-Castillo-RRcmdrv21.pdf> ## en castellano

• Libros:

Introductory Statistics with R. Peter Dalgaard, Springer-Verlag (2002).

Modern Applied Statistics with S, 4th Ed. by W.N. Venables and B.D. Ripley, Springer-Verlag (2002).

Libro libre: *Estadística Básica con R y R-Commander*. A. J. Arriaza Gómez, F. Fernández Palacín, M. A. López Sánchez, M. Muñoz Márquez, S. Pérez Plaza, A. Sánchez Navas. Servicio de Publicaciones de la Universidad de Cádiz (2008).
ISBN: 978-84-9828-186-6 **WWW:** <http://knuth.uca.es/ebrcmdr>

Fox, J. (2005) The R Commander: A Basic Statistics Graphical User Interface to R. *Journal of Statistical Software*, **14(9)**: 1–42.

Fox, J. (2007) Extending the R Commander by "plug in" packages. *R News*, **7(3)**: 46–52.

EVALUACIONES

Tres evaluaciones. Dos son parciales y una es global.

Las evaluaciones constan de dos partes:

- i) entrega de un trabajo,
- ii) preguntas teórico-prácticas.

El trabajo y las preguntas teórico prácticas son calificados en una escala de 1 a 10. Para aprobar la evaluación es necesario que estén aprobadas las dos partes (puntaje mayor a 6).

La nota de una evaluación aprobada resultará del promedio pesado de cada parte (0.6 para el trabajo y 0.4 para las preguntas teórico-prácticas).

Criterio de aprobación: Aprobar **una** evaluación parcial y la evaluación global.

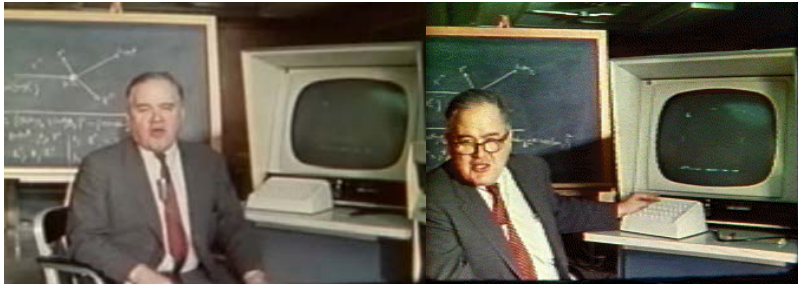
Las preguntas teórico-prácticas parciales no se recuperan.

Las preguntas teórico-prácticas correspondientes a la evaluación global pueden recuperarse una vez.

Si un trabajo no está aprobado, es devuelto para su corrección una vez.

¿Qué es el Análisis Exploratorio de Datos?

John W. Tukey (1977) introduce el Análisis Exploratorio de Datos aunque no presenta una definición precisa. Algunas de sus recomendaciones pueden darnos una idea¹:



“Exploratory data analysis . . . does not need probability, significance or confidence.”

“I hope that I have shown that exploratory data analysis is actively incisive rather than passively descriptive, with real emphasis on the discovery of the unexpected.”

“‘Exploratory data analysis’ is an attitude, a state of flexibility, a willingness to look for those things that we believe are not there, as well as those we believe to be there.”

“Exploratory data analysis isolates patterns and features of the data and reveals these forcefully to the analyst.”

“If we need a short suggestion of what exploratory data analysis is, I would suggest that: 1. it is an attitude, AND 2. a flexibility, AND 3. some graph paper (or transparencies, or both).”

¹ Jones, L. V. (Ed.). (1986). *The collected works of John W. Tukey: Philosophy and principles of data analysis 1949–1964* (Vols. III & IV). London: Chapman & Hall. en David R. Brillinger (2011) *Data Analysis, Exploratory* SAGE Publications. www.stat.berkeley.edu/~brill/Papers/EDASage.pdf

Tukey, J. W. (1962). The future of data analysis. *Annals of Mathematical Statistics*, 33, 1–67. http://www.stanford.edu/~gavish/documents/Tukey_the_future_of_data_analysis.pdf

“we need to stress flexibility of attack, willingness to iterate, and willingness to study things as they are, rather than as they, hopefully, should be”

¿Por qué R y R commander?

El R es un lenguaje de distribución libre que tiene

- un entorno flexible para el análisis de datos.
- una colección extensa y coherente de herramientas estadísticas para análisis de datos.
- un lenguaje para expresar modelos estadísticos y herramientas para utilizar modelos estadísticos lineales y no lineales.
- facilidades para el análisis de datos y su presentación tanto en la computadora como en papel.
- El R-Commander es una interfaz gráfica que facilita el uso de R sin utilizar los comandos directamente, pero generando los mismos.
- Tanto el R como el R-Commander (Rcmdr) se encuentran en permanente actualización.

Un poco de historia

R es una implementación libre, independiente, de fuente abierta (“open-source”) del lenguaje de programación S que actualmente es un producto comercial llamado S-PLUS.

El lenguaje S, que fue escrito a mediados de los años 70 en Bell Labs (de AT&T y actualmente Lucent Technologies).

Originalmente un programa para el sistema operativo Unix, **R** ahora puede obtenerse también en versiones para Windows y Macintosh y Linux.

El proyecto R fue iniciado por Robert Gentleman y Ross Ihaka (de donde se deriva “R”) del Statistics Department en la University of Auckland en 1995.

Actualmente R es mantenido por un grupo internacional de desarrolladores *voluntarios*: Core development team.

La página web del proyecto R es
<http://www.r-project.org/>.

R Commander es una interfaz gráfica (**graphical user interface GUI**) del lenguaje de programación R con licencia pública general “[GNU General Public License](#)”, desarrollado y mantenido por [John Fox](#) en el departamento de sociología en la [McMaster University](#).

Ver *Estadística Básica con R y R-Commander* (2008) para más detalles sobre este tipo de licencia

El interfaz, en informática, es un elemento de [conexión](#) que facilita el intercambio de datos, como por ejemplo el [teclado](#), un tipo de interfaz entre el usuario y la computadora.

La pagina web del R Commander es

<http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/>

Encontramos una descripción del menú del programa

R Commander Menu Tree [current as of version 1.9-0]

```
File - Change working directory
      | - Open script file
      | - Save script
      | - Save script as
      | - Save output
      | - Save output as
      | - Save R workspace
      | - Save R workspace as
      | - Exit - from Commander
      |   | - from Commander and R
Edit - Cut
      | - Copy
      | - Paste
      | - Delete
      | - Find
      | - Select all
      | - Undo
      | - Redo
      | - Clear Window
Data - New data set
      | - Load data set
      | - Merge data sets
      | - Import data - from text file, clipboard, or URL
      |   | - from SPSS data set
      |   | - from SAS xport file
      |   | - from Minitab data set
      |   | - from STATA data set
      |   | - from Excel, Access, or dBase data set [32-bit Windows only]
      |   | - from Excel file [currently 64-bit Windows only]
      | - Data in packages - List data sets in packages
      |   | - Read data set from attached package
      | - Active data set - Select active data set
      |   | - Refresh active data set
      |   | - Help on active data set (if available)
      |   | - Variables in active data set
      |   | - Set case names
      |   | - Subset active data set
      |   | - Aggregate variables in active data set
      |   | - Remove row(s) from active data set
      |   | - Stack variables in active data set
      |   | - Remove cases with missing data
      |   | - Save active data set
      |   | - Export active data set
      | - Manage variables in active data set - Recode variable
      |   | - Compute new variable
      |   | - Add observation numbers to data set
      |   | - Standardize variables
      |   | - Convert numeric variables to factors
      |   | - Bin numeric variable
      |   | - Reorder factor levels
      |   | - Define contrasts for a factor
      |   | - Rename variables
      |   | - Delete variables from data set
Statistics - Summaries - Active data set
      |   | - Numerical summaries
      |   | - Frequency distributions
      |   | - Count missing observations
      |   | - Table of statistics
      |   | - Correlation matrix
      |   | - Correlation test
      |   | - Shapiro-Wilk test of normality
      | - Contingency Tables - Two-way table
      |   | - Multi-way table
      |   | - Enter and analyze two-way table
```

```

|- Means - Single-sample t-test
|         |- Independent-samples t-test
|         |- Paired t-test
|         |- One-way ANOVA
|         |- Multi-way ANOVA
|- Proportions - Single-sample proportion test
|         |- Two-sample proportions test
|- Variances - Two-variances F-test
|         |- Bartlett's test
|         |- Levene's test
|- Nonparametric tests - Two-sample Wilcoxon test
|         |- Paired-samples Wilcoxon test
|         |- Kruskal-Wallis test
|         |- Friedman rank-sum test
|- Dimensional analysis - Scale reliability
|         |- Principal-components analysis
|         |- Factor analysis
|         |- Confirmatory factor analysis
|         |- Cluster analysis - k-means cluster analysis
|         |- Hierarchical cluster analysis
|         |- Summarize hierarchical clustering
|         |- Add hierarchical clustering to data set
|- Fit models - Linear regression
|         |- Linear model
|         |- Generalized linear model
|         |- Multinomial logit model
|         |- Ordinal regression model

Graphs - Color palette
|       |- Index plot
|       |- Histogram
|       |- Stem-and-leaf display
|       |- Boxplot
|       |- Quantile-comparison plot
|       |- Scatterplot
|       |- Scatterplot matrix
|       |- Line graph
|       |- XY conditioning plot
|       |- Plot of means
|       |- Strip chart
|       |- Bar graph
|       |- Pie chart
|       |- 3D graph - 3D scatterplot
|       |         |- Identify observations with mouse
|       |         |- Save graph to file
|       |- Save graph to file - as bitmap
|       |         |- as PDF/Postscript/EPS
|       |         |- 3D RGL graph

Models - Select active model
|       |- Summarize model
|       |- Add observation statistics to data
|       |- Confidence intervals
|       |- Akaike Information Criterion (AIC)
|       |- Bayesian Information Criterion (BIC)
|       |- Stepwise model selection
|       |- Subset model selection
|       |- Hypothesis tests - ANOVA table
|       |         |- Compare two models
|       |         |- Linear hypothesis
|       |- Numerical diagnostics - Variance-inflation factors
|       |         |- Breusch-Pagan test for heteroscedasticity
|       |         |- Durbin-Watson test for autocorrelation
|       |         |- RESET test for nonlinearity
|       |         |- Bonferroni outlier test
|- Graphs - Basic diagnostic plots
|         |- Residual quantile-comparison plot
|         |- Component+residual plots
|         |- Added-variable plots
|         |- Influence plot
|         |- Effect plots

Distributions - Continuous distributions - Normal distribution - Normal quantiles
|         |         |- Normal probabilities
|         |         |- Plot normal distribution
|         |         |- Sample from normal distribution
|         |         |- t distribution - t quantiles
|         |         |         |- t probabilities
|         |         |         |- Plot t distribution
|         |         |         |- Sample from t distribution
|         |         |- Chi-squared distribution - Chi-squared quantiles
|         |         |         |- Chi-squared probabilities
|         |         |         |- Plot chi-squared distribution

```

```

|         |         |- Sample from chi-squared distribution
|- F distribution - F quantiles
|         |         |- F probabilities
|         |         |- Plot F distribution
|         |         |- Sample from F distribution
|- Exponential distribution - Exponential quantiles
|         |         |- Exponential probabilities
|         |         |- Plot exponential distribution
|         |         |- Sample from exponential distribution
|- Uniform distribution - Uniform quantiles
|         |         |- Uniform probabilities
|         |         |- Plot uniform distribution
|         |         |- Sample from uniform distribution
|- Beta distribution - Beta quantiles
|         |         |- Beta probabilities
|         |         |- Plot beta distribution
|         |         |- Sample from beta distribution
|- Cauchy distribution - Cauchy quantiles
|         |         |- Cauchy probabilities
|         |         |- Plot Cauchy distribution
|         |         |- Sample from Cauchy distribution
|- Logistic distribution - Logistic quantiles
|         |         |- Logistic probabilities
|         |         |- Plot logistic distribution
|         |         |- Sample from logistic distribution
|- Lognormal distribution - Lognormal quantiles
|         |         |- Lognormal probabilities
|         |         |- Plot lognormal distribution
|         |         |- Sample from lognormal distribution
|- Gamma distribution - Gamma quantiles
|         |         |- Gamma probabilities
|         |         |- Plot gamma distribution
|         |         |- Sample from gamma distribution
|- Weibull distribution - Weibull quantiles
|         |         |- Weibull probabilities
|         |         |- Sample from Weibull distribution
|- Gumbel distribution - Gumbel quantiles
|         |         |- Gumbel probabilities
|         |         |- Plot Gumbel distribution
|         |         |- Sample from Gumbel distribution
|- Discrete distributions - Binomial distribution - Binomial quantiles
|         |         |         |- Binomial tail probabilities
|         |         |         |- Binomial probabilities
|         |         |         |- Plot binomial distribution
|         |         |         |- Sample from binomial distribution
|- Poisson distribution - Poisson quantiles
|         |         |         |- Poisson tail probabilities
|         |         |         |- Poisson probabilities
|         |         |         |- Plot Poisson distribution
|         |         |         |- Sample from Poisson distribution
|- Geometric distribution - Geometric quantiles
|         |         |         |- Geometric tail probabilities
|         |         |         |- Geometric probabilities
|         |         |         |- Plot geometric distribution
|         |         |         |- Sample from geometric distribution
|- Hypergeometric distribution - Hypergeometric quantiles
|         |         |         |- Hypergeometric tail probabilities
|         |         |         |- Hypergeometric probabilities
|         |         |         |- Plot hypergeometric distribution
|         |         |         |- Sample from hypergeometric distribution
|- Negative binomial distribution - Negative binomial quantiles
|         |         |         |- Negative binomial tail probabilities
|         |         |         |- Negative binomial probabilities
|         |         |         |- Plot negative binomial distribution
|         |         |         |- Sample from negative binomial distr.

Tools - Load package(s)
|       |- Load Rcmdr plug-in(s)
|       |- Options
Help - Commander help
|       |- Introduction to the R Commander
|       |- Help on active data set (if available)
|       |- About Rcmdr
|       |- Start R help system

```

A continuación veremos cómo bajar, instalar y comenzar a usar el R y su

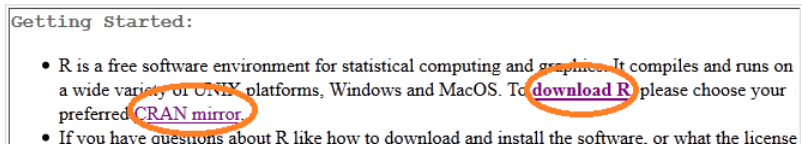
interfaz gráfica R-Commander. Primero se debe instalar el R y luego se le "incorpora" el R-Commander.

Instalación de R

Las ventanas que se presentan a continuación pueden presentar diferencias.

El sitio principal de R es

<http://www.r-project.org/>

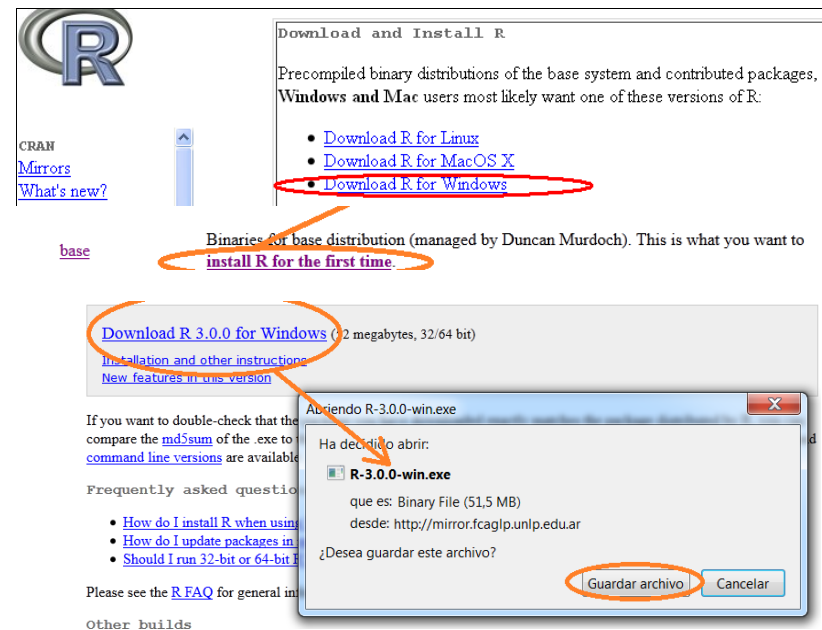


Clickeando en CRAN mirror se accede al sitio más cercano

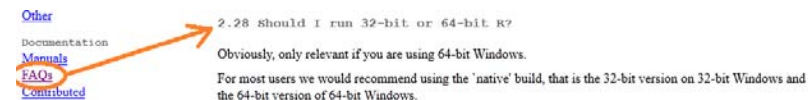


Seleccionamos

<http://mirror.fcaglp.unlp.edu.ar/CRAN/>

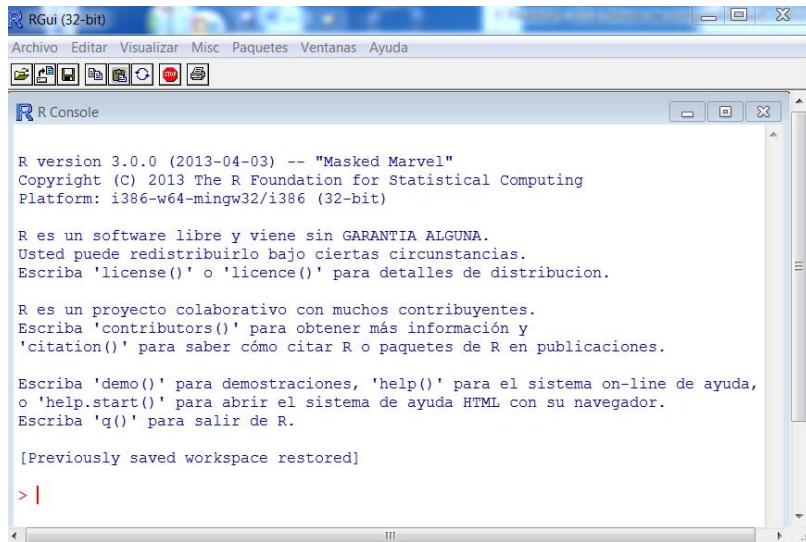


Ejecutar el archivo .exe para instalar el R



Iniciando R

Una forma de usar R es en forma interactiva mediante la línea de comandos. Una vez instalado hay que hacer un doble click en el ícono de R (en Unix/Linux, se escribe R desde el símbolo de comandos (command prompt)). Cuando R se inicia, aparece la ventana del programa "Gui" (graphical user interface) con un mensaje de apertura



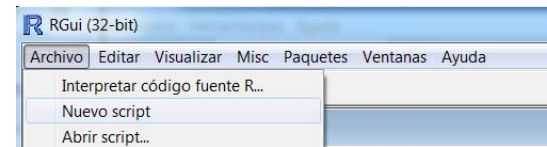
Debajo del mensaje de apertura en la Consola de R se encuentra el “prompt” que es el símbolo `>` (“mayor”).

`>`

La mayoría de las expresiones en R se escriben directamente a continuación del “prompt” en la Consola de R.

También es posible generar un **archivo “script”** con los comandos que se quieren ejecutar.

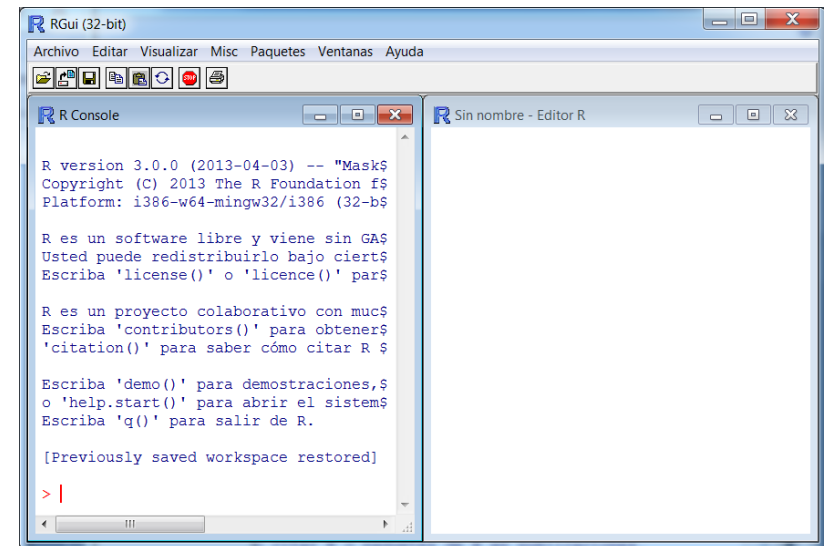
Ventana de Escritura (Script)



Archivo -> Nuevo script

Es útil acomodar las dos ventanas, la de comandos y la del editor, para poder verlas simultáneamente

Ventanas -> Divida Verticalmente



Conceptos básicos del lenguaje de R

expresiones, asignaciones, funciones, tipos de datos

expresión simple

```
> 2+3 ↵ (enter) #expresión
[1] 5      #evaluación
```

expresión un poco más compleja

```
> sqrt(3/4)/(1/3 - 2/pi^2)
[1] 6.626513
```

El símbolo > (prompt) indica la línea de comandos y el [1] que la respuesta comienza en el primer elemento de un vector.

Si se escribe una expresión incompleta > 2* ↵
R responde con un + que indica continuar

```
> 2* ↵
+ 5
[1] 10
```

```
> sqrt(3/4 ↵
+ )
[1] 0.8660254
```

La expresión más común es el llamado a una función, se escribe el nombre de la función seguida de sus argumentos entre paréntesis.

```
> sqrt(3/4)
[1] 0.8660254
```

```
> Sqrt(3/4)
Error: no se pudo encontrar la función "Sqrt"
> sqrt(3/4)
[1] 0.8660254
```

Diferencia mayúsculas de minúsculas.

El software “reconoce” al número pi

```
> pi
[1] 3.141593
```

Si se escribe una cadena de caracteres seguidos por un par de paréntesis R lo interpreta como el **nombre** de una **función**.

```
> pi()
```

```
Error: no se pudo encontrar la función
"pi"
```

Si la función existe y no requiere argumentos se ejecutará la función, como

```
q()
```

para irse de la sesión.

Si la función existe y requiere argumentos dará un mensaje de error

```
> sqrt()
Error en sqrt() : 0 arguments passed to 'sqrt' which requires 1
```

Algunos mensajes de error aparecen en Castellano y otros en Inglés!

Asignaciones

Hay varios operadores con los que es posible realizar asignaciones

"<-" signo menor seguido del signo menos, sin espacios

"=",

Enteros consecutivos

```
> a <- 2:6      # crea el vector (2,3,4,5,6)
```

```
> a
[1] 2 3 4 5 6
```

Aritmética

```
> b <- 2*a+1
> b
[1] 5 7 9 11 13
```

```
> b <- a/2      # división
> b
[1] 1.0 1.5 2.0 2.5 3.0
```

```
> b <- a^3.7    # eleva a la potencia 3.7 cada
componente de a
```

```
> b
[1] 12.99604 58.25707 168.89701
```

```
[4] 385.64616 757.11112
```

```
> b <- log(a)    # logaritmo natural
> b
[1] 0.6931472 1.0986123 1.3862944
[4] 1.6094379 1.7917595
```

```
> b <- log10(a)  # asignación
> log10(a)      # evaluación
[1] 0.3010300 0.4771213 0.6020600
[4] 0.6989700 0.7781513
> b <- logb(a,base=2) # logaritmo base 2
```

```
> b <- logb(a,2) # idem
```

```
> help(logb)    # se abre una ventana de
#ayuda
```

Instalar el Rcmdr

Para instalar el Rcmdr:

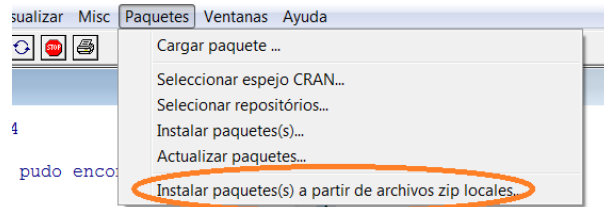
desde el R ir a Paquetes→ Seleccionar espejo CRAN (chooseCRANmirror) -> (La Plata)
Instalar Paquete(s) y el mirror desde el cual se quiere instalar el paquete (por ej. el de la
Universidad Nacional de La Plata). Luego se selecciona Rcmdr.

Si algo falla, otra forma es:

desde la página <http://www.r-project.org/>
elegir el espejo (Universidad Nacional de La Plata)
ir a "Packages" -> "Table of available packages, sorted by name"
seleccionar Rcmdr -> Windows binary Rcmdrzip

Guardar archivo.

Desde R



Paquetes → Instalar paquetes a partir de .zip locales

Si esto tampoco anda: descomprimir la carpeta .zip y copiarla en la carpeta "library" de instalación del R (por ejemplo C:\Program Files\R\R-3.0.0\library)

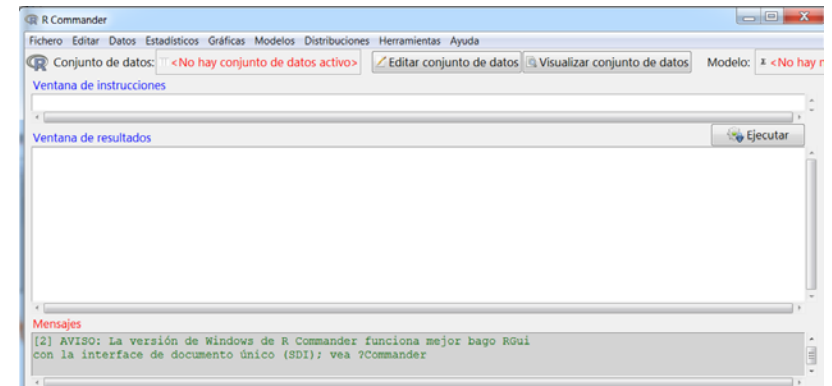
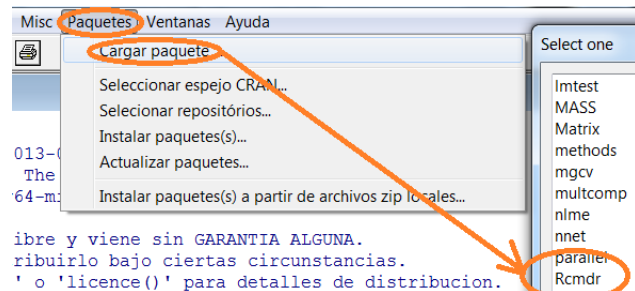
También bajar los paquetes adicionales: copiarlos a una carpeta y luego
Paquetes → Instalar paquetes a partir de .zip locales

El R Commander se inicia como toda biblioteca de R

Opción 1: Desde la consola de R, ejecutar `library("Rcmdr")`

```
> library("Rcmdr")
```

Opción 2: Desde el menú, Paquetes -> Cargar Paquete -> Rcmdr



Seguimos con R

Tipos de datos

5 Tipos de objetos datos básicos:

data frames, matrices, vectores, listas y funciones.

Data frame (marco de datos)

Ejemplo de un conjunto de datos incorporado en el R

```
> library(rpart)
> kyphosis
  Kyphosis Age Number Start
1   absent  71      3     5
2   absent 158      3    14
3   present 128      4     5
4   absent   2      5     1
5   absent   1      4    15
6   absent   1      2    16
7   absent  61      2    17
8   absent  37      3    16
9   .....
```

en cada columna se guardan los valores de una variable.

```
> class(kyphosis)
[1] "data.frame"
> ?kyphosis
```

data frame, 81 filas (casos, niños sometidos a una cirugía espinal) Variables: Kyphosis - factor - con 2 niveles presencia o ausencia de una deformidad post operatoria, Age, Number, Start son vectores numéricos.

Todas las columnas de un data frame deben tener la misma longitud: intentamos generar un data frame con columnas de diferente longitud

```
> data.frame( x=1:4,y=1:3)
Error in data.frame(x = 1:4, y = 1:3) :
  arguments imply differing number of rows: 4, 3
```

Corregimos el error

```
> data.frame( x=1:3,y=1:3)
  x y
1 1 1
2 2 2
3 3 3
```

¿Pero si necesitamos variables con diferente longitud? Corregimos el error de otra forma:

```
> data.frame( x=1:4,y=c(1:3,NA))
  x y
1 1 1
2 2 2
3 3 3
4 4 NA
```

¿Cual es el resultado de aplicar la función c?

¿Que significa "NA"?

Para ver todos los datos incorporados en R:

```
> data(package = .packages(all.available =  
TRUE))
```

Aparecen unos mensajes de advertencias (warnings) que pasamos por alto y se abre una ventana con el listado de todos los conjuntos de datos incorporados al R

Veremos también:

Matrices: son similares a los data frames, salvo que sus elementos deben tener datos con el mismo modo (carácter, numérico, lógico). Las filas y las columnas pueden tener nombres.

Vectores: es un conjunto ordenado de elementos que tienen el mismo modo. Los elementos de un vector pueden tener nombres.

Listas: son colecciones de otros objetos. Sus componentes pueden ser data frames, matrices, vectores, otras listas, cualquier objeto.

Funciones: existen gran cantidad de funciones incorporadas. También es posible agregar funciones definidas por el usuario.