

R y RStudio



2022

Inicios del R

- En 1976 S John Chambers old de Bell Telephone Laboratories, inician el programa S
- En 1993 se vende S y se crea el comercial (**S-plus**)
- R se considera la **versión libre del programa S**, de los Laboratorios Bell.
- En 1993 Ross Ihaka y Robert Gentleman crean la primera versión de R.
- En 1995, se crea el GNU General Public License para hacer libre al R.
- En 1997 se crea el R Core Group

R es un software libre

- R es software libre.
- Los derechos de autor del código fuente principal de R pertenecen a la Fundación R y se publican bajo la licencia pública general GNU.
 1. La libertad de ejecutar el programa, para cualquier propósito.
 2. La libertad de estudiar cómo funciona el programa y adaptarlo a sus necesidades.
 3. La libertad de redistribuir copias.
 4. La libertad de mejorar el programa y divulgar sus mejoras al público, de modo que toda la comunidad se beneficie.

¿Que es R-UCA

- R-UCA para Windows es una instalación única de **R** junto a **R-commander** y a algunos **paquetes de uso frecuente**.

- **Ventajas :**

- Se instala en un solo paso, con un solo archivo y un solo click.
- Su instalación no requiere internet.
- Rcmdr inicia automáticamente
- Se instala en forma autónoma y se desinstala sin problemas.

¿Que es RStudio?

- **RStudio** es una interface que facilita el uso de el R y expande sus capacidades
- Se puede descargar en la siguiente dirección:

<https://www.rstudio.com/products/rstudio/download/>



Robert Gentleman (izquierda) y Ross Ihaka (derecha) creadores de R

Software IEEE Spectrum ranking

Software	Rank
Scala	63.5
Julia	56
Rust	55.6
Shell	52
Processing	49.2
C#	48.1
SAS	45.2
Fortran	43
Cuda	41
Visual Basic	40.3
Delphi	38.6
Perl	38.2

Ranking

The Top 14 Statistical Analysis Software for Enterprise

Check out this list of the top Statistical Analysis Software products based on user satisfaction. A product's satisfaction score is calculated by a [proprietary algorithm](#) that factors in real-user satisfaction ratings from review data. Software buyers can compare products according to their satisfaction scores to streamline the buying process and quickly identify the best products based on the experiences of their peers.



Filters

Enterprise (>1000 emp.)

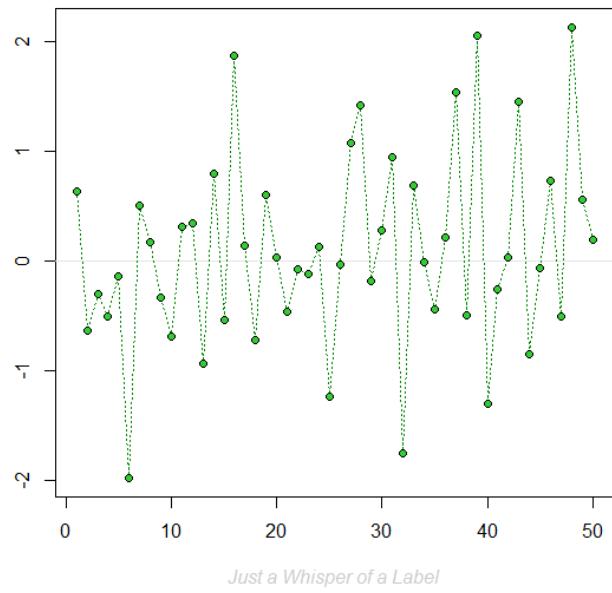
All Regions

Collapse All ↗	G2 Satisfaction Score ⓘ	Compare
▼ #1 RStudio	93	<input checked="" type="checkbox"/> Compare
▼ #2 OriginPro	86	<input type="checkbox"/> Compare
▼ #3 TIMI Suite	85	<input type="checkbox"/> Compare
▼ #4 IBM SPSS Statistic...	84	<input type="checkbox"/> Compare
▼ #5 JMP	75	<input type="checkbox"/> Compare
▼ #6 Stata	75	<input type="checkbox"/> Compare
▼ #7 Minitab	57	<input checked="" type="checkbox"/> Compare
▼ #8 Grapher	55	<input type="checkbox"/> Compare
▼ #9 GraphPad Prism	38	<input type="checkbox"/> Compare
▼ #10 KNIME Analytics...	33	<input type="checkbox"/> Compare
▼ #11 eviews	31	<input type="checkbox"/> Compare
▼ #12 SAS/STAT	20	<input type="checkbox"/> Compare

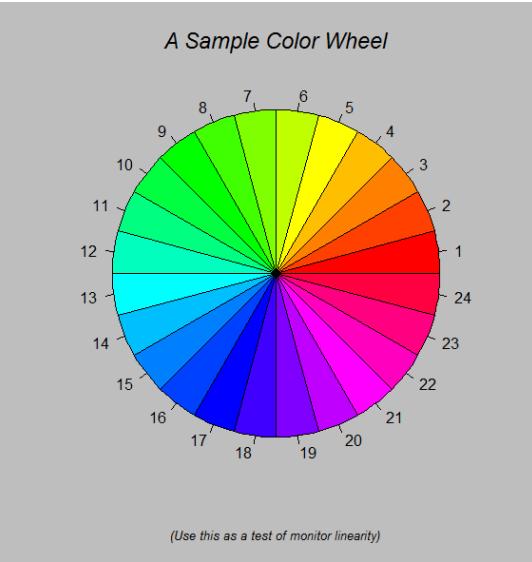
Ventajas del R

- Es muy poderoso
- Es gratis y multiplataforma (Windows, linux, Mac)
- Gráficos de alta calidad
- R hace que sea fácil escribir un programa para procesar y analizar datos
- Lenguaje ~~orientado a objetos con sintaxis (relativamente) intuitiva.~~

Simple Use of Color In a Plot

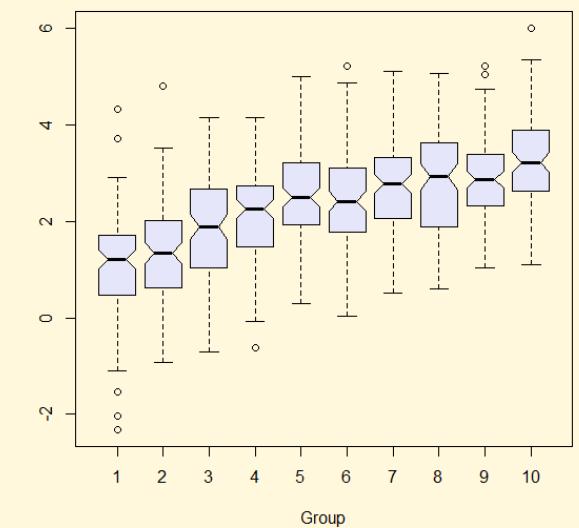


A Sample Color Wheel

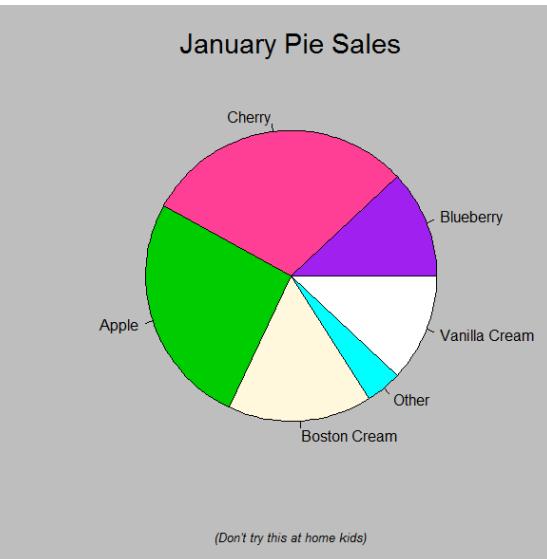


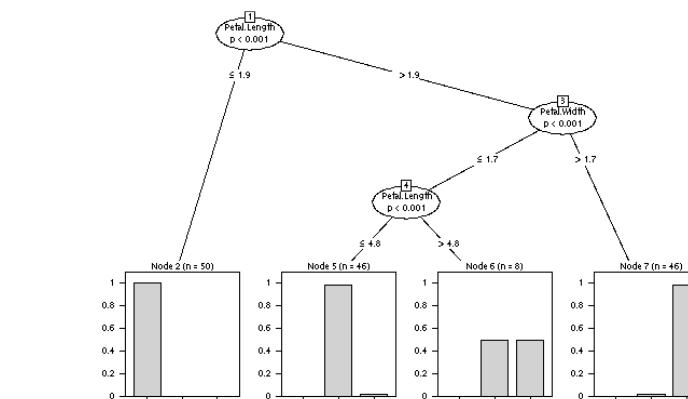
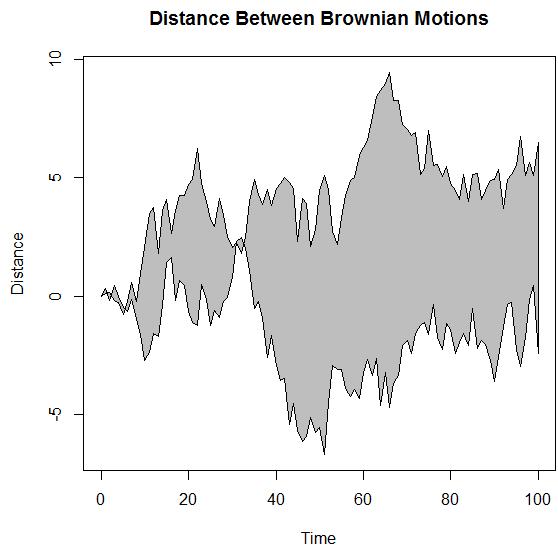
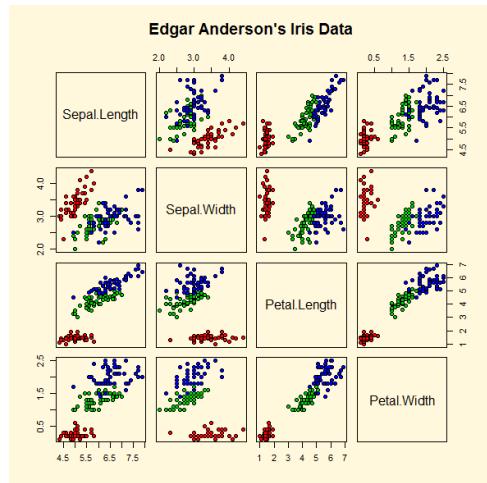
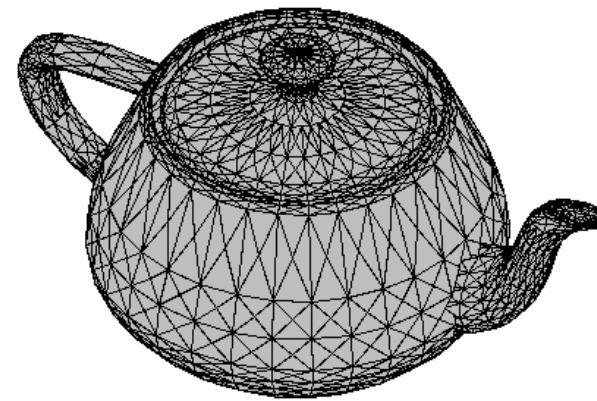
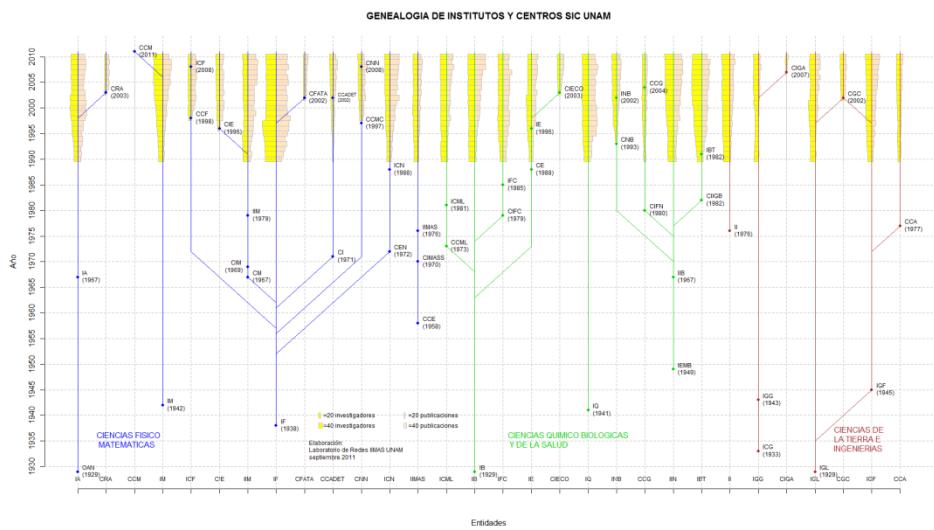
Gráficos en R

Notched Boxplots

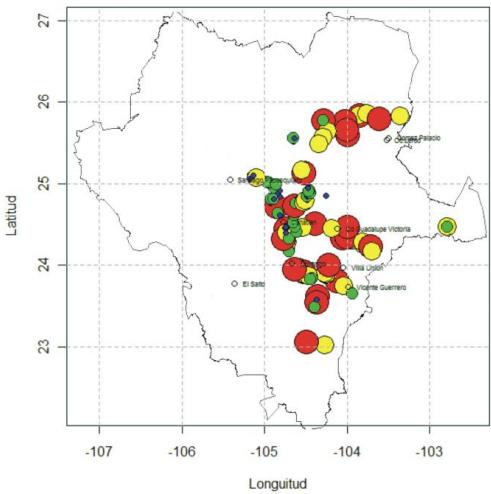


January Pie Sales

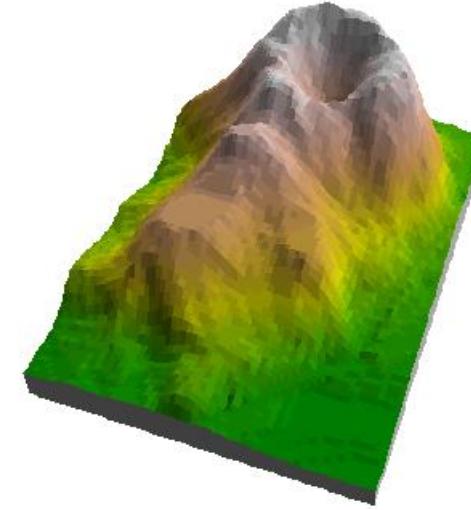
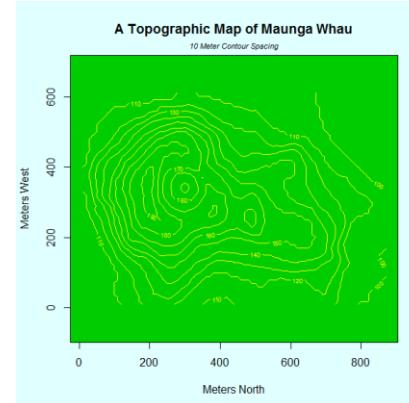




DURANGO RENDIMIENTO



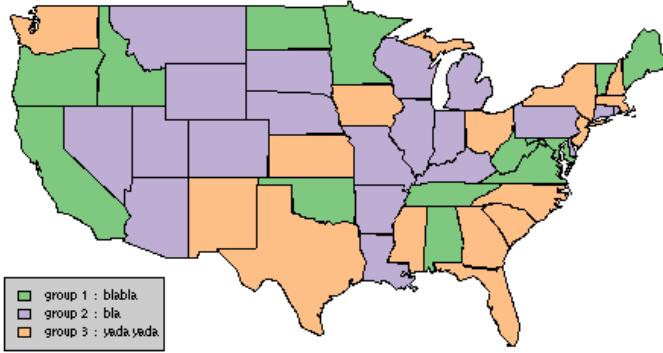
A Topographic Map of Maunga Whau



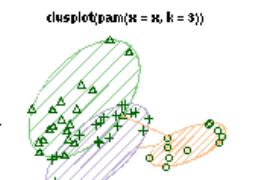
Gráficos en R

Cluster analysis on state x 77 data

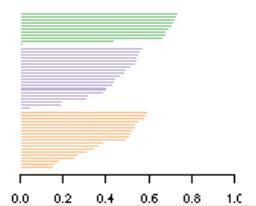
variables : Illiteracy, Life Exp, Murder, HS Grad



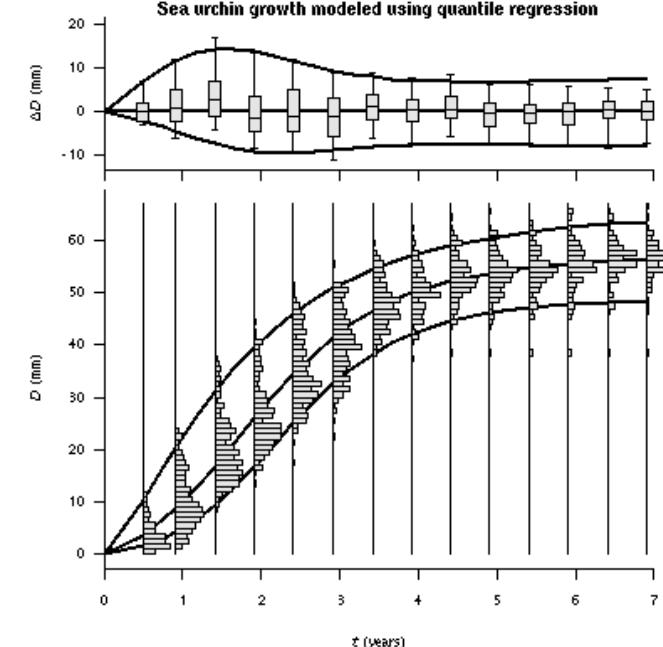
clusplot(pam(x = x, k = 3))

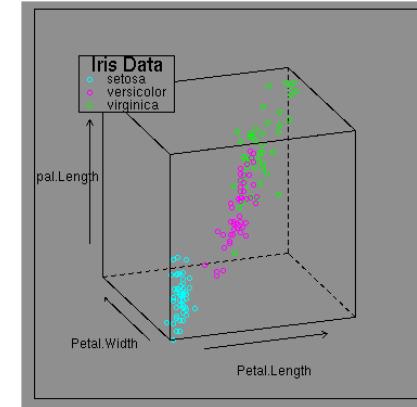
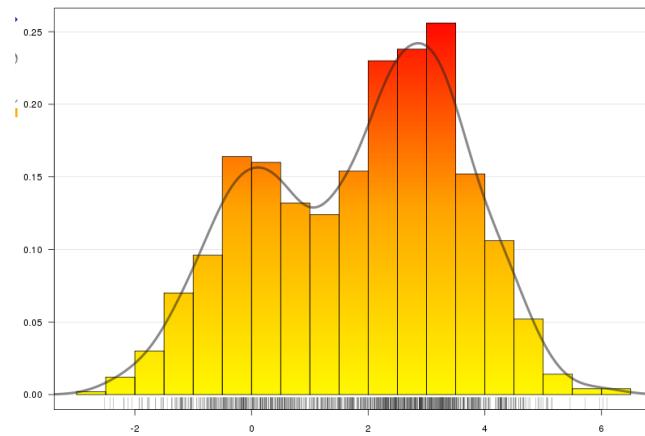


Silhouette plot of pam(x = x, k = 3)

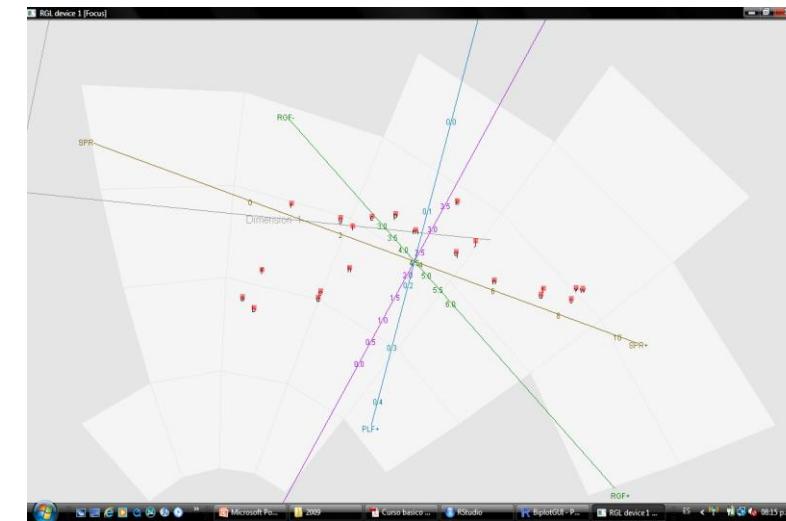
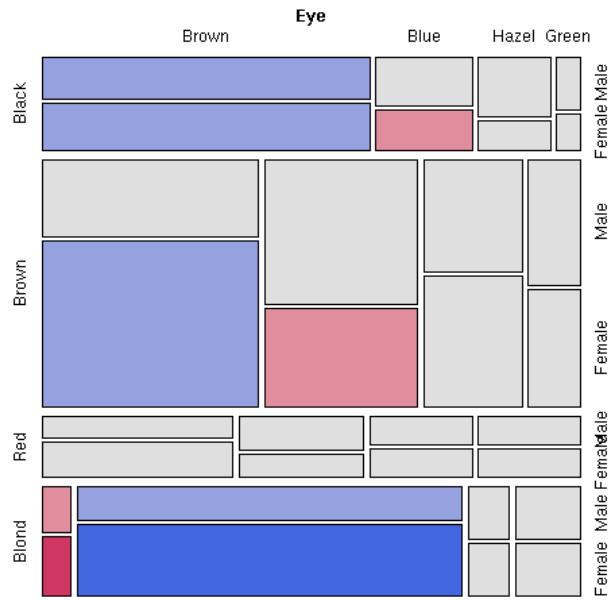


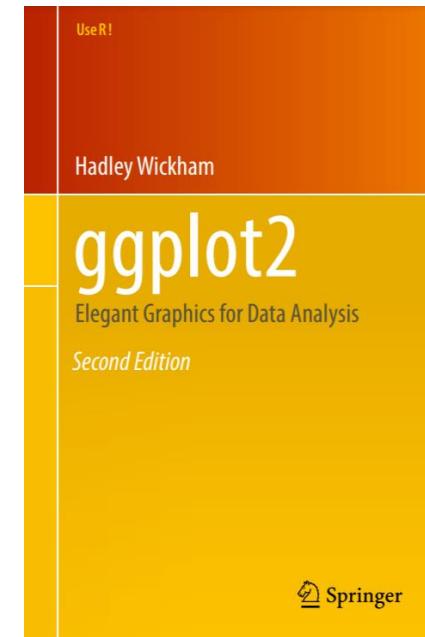
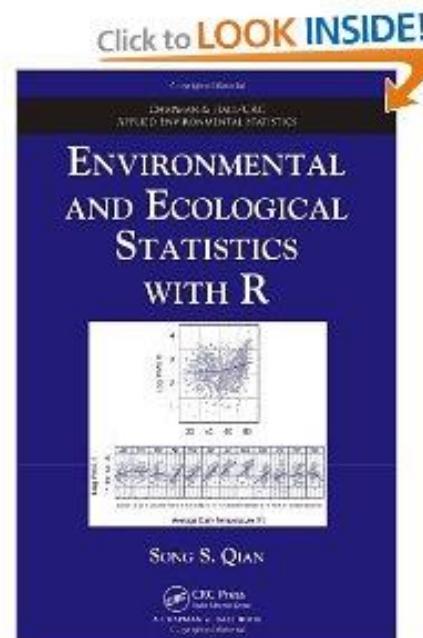
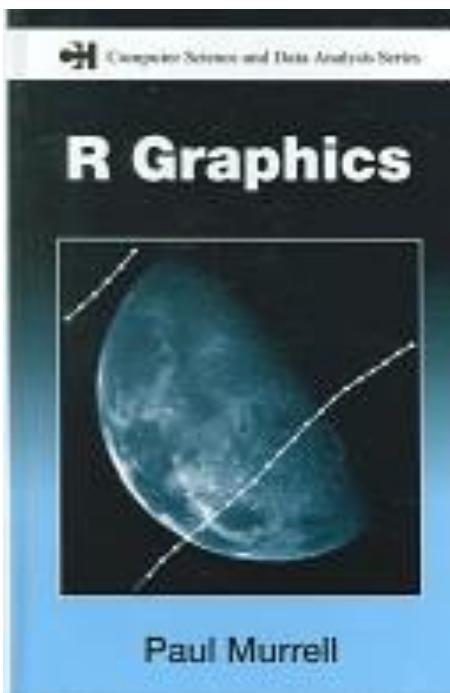
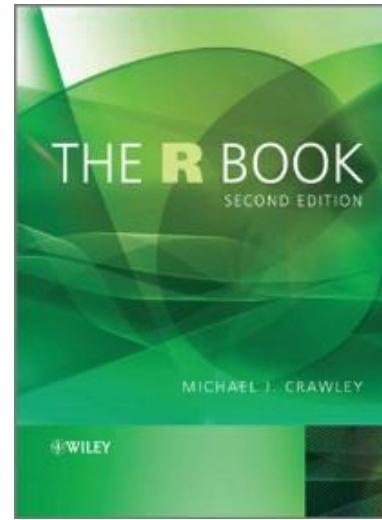
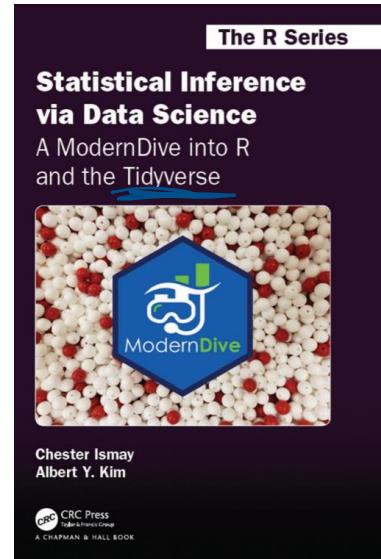
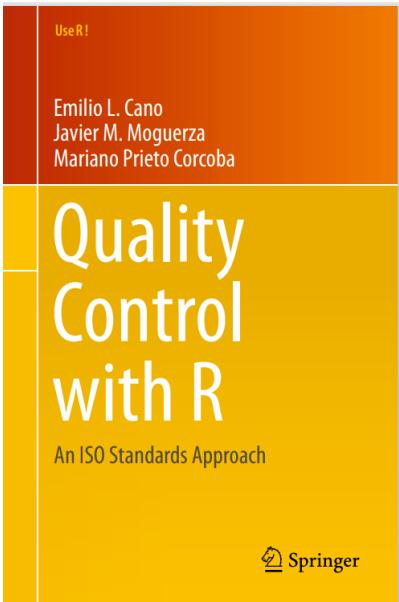
Sea urchin growth modeled using quantile regression





Gráficos en R





Desventajas del R

- No existe soporte técnico.
- No existe una interfaz “amigable”.
- Requiere de precisión en la sintaxis.

Instalación de R y RStudio

- Descargar las últimas versiones de R y RStudio.
- Aunque probablemente usará RStudio como su consola principal y editor, primero debe instalar.

Orden de instalación:

- **primero instale R y**
- luego RStudio.**

Proyecto R-UCA

http://knuth.uca.es/R/doku.php?id=instalacion_de_r_y_rcmdr:r-uca

- R-UCA para Windows es una instalación única de **R** junto a **R-commander** y a algunos **paquetes de uso frecuente**.

• **Ventajas :**

- Se instala en un solo paso, con un solo archivo y un solo click.
- Su instalación no requiere internet.
- Rcmdr inicia automáticamente
- Se instala en forma autónoma y se desinstala sin problemas.

Conceptos Importantes

- R ES UN LENGUAJE DE PROGRAMACIÓN
- Rstudio es una interfaz
- Se basa en funciones
 - Función (argumentos)
 - Ejemplo: `mean(datos$X)`

Conceptos Importantes

- Una Fortaleza importante son los paquetes
 - Paquetes = rutinas generados por los usuarios
 - Ejemplos: `ggplot2` y `nmlm`
- Los paquetes se instalan una sola vez
 - Ejemplo: `install.packages ("ggplot2")`
- Los paquetes deben activarse cuando se van a usar
 - Solo una vez por sesión
 - ejemplo: `library(ggplot2)`

Conceptos Importantes

- Se puede trabajar en la consola o en un “script”
- “Script” es un programa que puede guardarse
 - Se puede modificar fácilmente
 - Ayuda a la reproductibilidad de los resultados
 - Como todo programa puede ser “corrido” por línea o completo
- Los resultados pueden ser almacenado en un objeto y visualizarse en cualquier momento

Para descargar R:



[CRAN](#)
[Mirrors](#)
[What's new?](#)
[Task Views](#)
[Search](#)

[About R](#)
[R Homepage](#)
[The R Journal](#)

[Software](#)
[R Sources](#)
[R Binaries](#)
[Packages](#)
[Other](#)

[Documentation](#)
[Manuals](#)
[FAQs](#)
[Contributed](#)

R-4.1.1 for Windows (32/64 bit)

[Download R 4.1.1 for Windows](#) (86 megabytes, 32/64 bit)

[Installation and other instructions](#)
[New features in this version](#)

If you want to double-check that the package you have downloaded matches the package distributed by CRAN, you can compare the [md5sum](#) of the .exe to the [fingerprint](#) on the master server. You will need a version of md5sum for windows: both [graphical](#) and [command line versions](#) are available.

Frequently asked questions

- [Does R run under my version of Windows?](#)
- [How do I update packages in my previous version of R?](#)
- [Should I run 32-bit or 64-bit R?](#)

Please see the [R FAQ](#) for general information about R and the [R Windows FAQ](#) for Windows-specific information.

Other builds

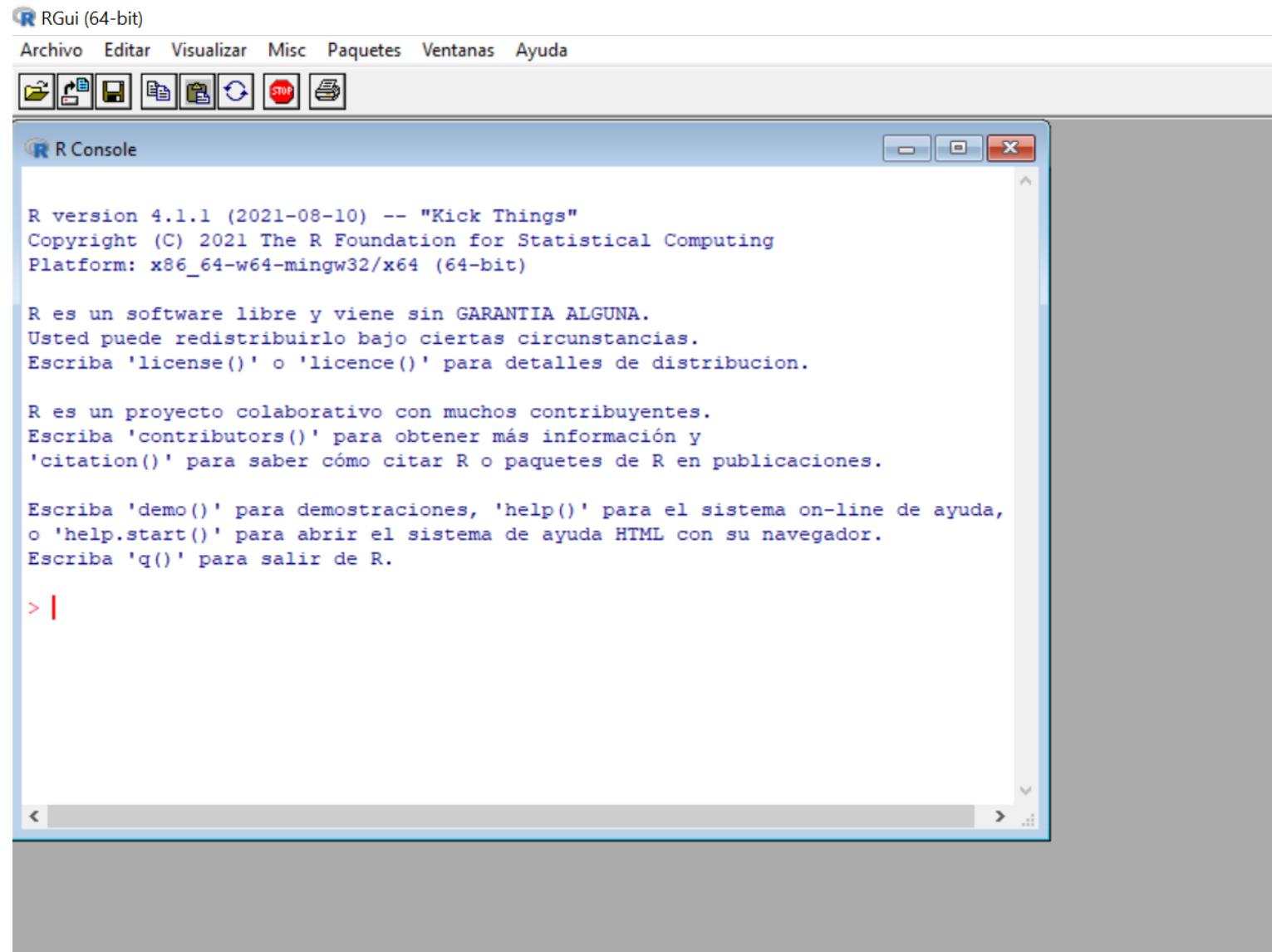
- Patches to this release are incorporated in the [r-patched snapshot build](#).
- A build of the development version (which will eventually become the next major release of R) is available in the [r-devel snapshot build](#).
- [Previous releases](#)

Note to webmasters: A stable link which will redirect to the current Windows binary release is
<CRAN MIRROR>/bin/windows/base/release.html

Para Instalar “R”

- 1. Abra la carpeta donde guardó el archivo .exe;
- 2. Haga doble clic en el archivo R-X.x.x-win.exe;
- 3. Acepte las opciones predeterminadas en el asistente de instalación.
- [\(205\) Installing R on Mac OSX - YouTube](#)

Iniciando consola R



RStudio - un IDE para R

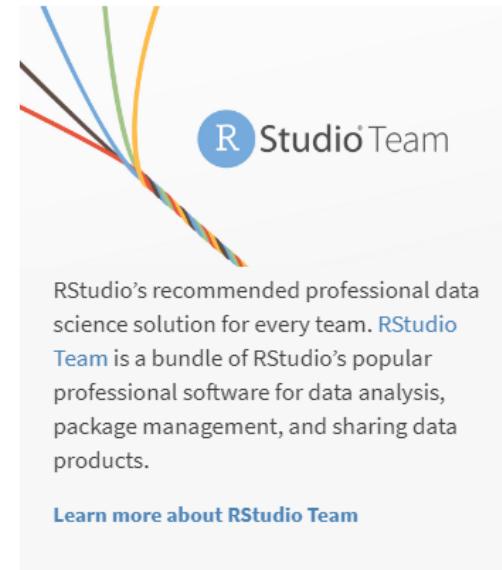
- Aunque podemos usar R directamente, es recomendable instalar y usar un entorno integrado de desarrollo (IDE, por sus siglas en inglés).
- RStudio ha contribuido de manera significativa para lograr que R sea lenguaje de programación más accesible.
- Si ya hemos instalado R en nuestro equipo, RStudio lo detectará automáticamente.

Para descargar RStudio:

Choose Your Version

The RStudio IDE is a set of integrated tools designed to help you be more productive with R and Python. It includes a console, syntax-highlighting editor that supports direct code execution, and a variety of robust tools for plotting, viewing history, debugging and managing your workspace.

[LEARN MORE ABOUT THE RSTUDIO IDE](#)



RStudio's recommended professional data science solution for every team. [RStudio Team](#) is a bundle of RStudio's popular professional software for data analysis, package management, and sharing data products.

[Learn more about RStudio Team](#)

RStudio Desktop

Open Source License

Free

RStudio Desktop Pro

Commercial License

\$995

/year

RStudio Server

Open Source License

Free

RStudio Workbench

•

Commercial License

\$4,975

/year

(5 Named Users)

[DOWNLOAD](#)

[Learn more](#)

[BUY](#)

[Learn more](#)

[DOWNLOAD](#)

[Learn more](#)

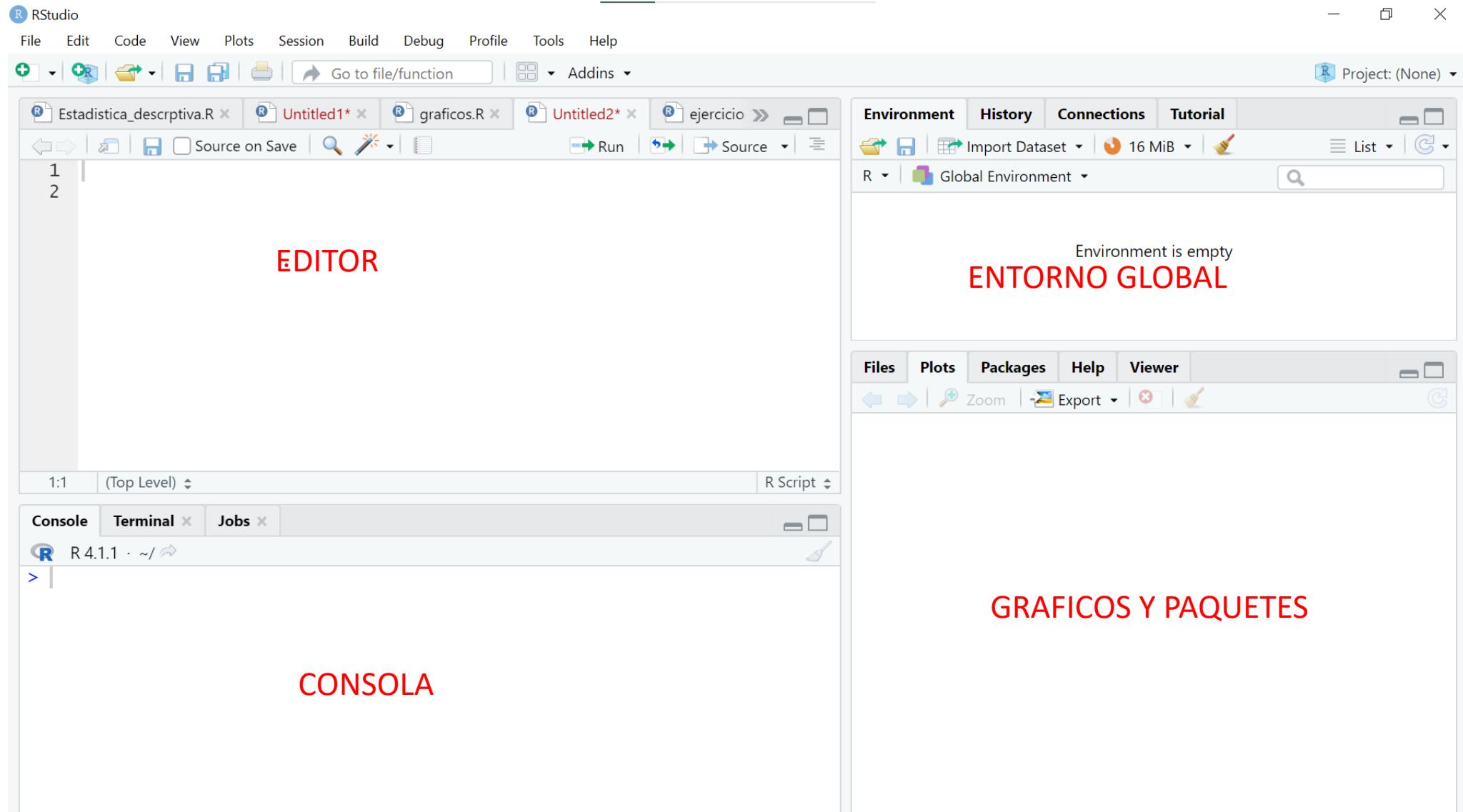
[BUY](#)

[Evaluation | Learn more](#)

Para instalar RStudio

1. Asegúrese de tener la última versión de Java instalada en su sistema;
2. Abra la carpeta donde guardó el archivo **.exe**;
3. Haga doble clic en el archivo **RStudio-X.xx.xxx.exe**;
4. Acepte las opciones predeterminadas en el asistente de instalación.

Pantalla inicial en Rstudio



Interfaz RStudio

- La interfaz RStudio tiene cuatro paneles principales:
 - **Consola**: donde puede escribir comandos y ver la salida. La consola es todo lo que vería si ejecutara R en la línea de comando sin RStudio.
 - **Editor de scripts**: donde puede escribir comandos y guardarlos en un archivo. También puede enviar los comandos para que se ejecuten en la consola.
 - **Entorno / Historial**: el entorno muestra todos los objetos activos y el historial realiza un seguimiento de todos los comandos que se ejecutan en la consola
 - **Archivos / Plots / Paquetes / Ayuda** RStudio

Modos de interacción con R

- De forma interactiva:
 - Escribe comandos en la consola R línea por línea y obtenga comentarios directos
- A través de archivos de secuencia de comandos (script):
 - Escribe comandos en un archivo de texto como una secuencia de comandos y luego puede:
 - Copiar y pegar comandos en la consola
 - Leer y ejecutar todos los comandos a la vez

Preferible usar archivos de script

- Promueve:
 - Programación / análisis organizados
 - Reutilización de código
 - Aumenta la replicabilidad
 - Más fácil de corregir errores
- Puede escribir / editar archivos de script con:
 - El editor incorporado
 - Un editor externo
 - Un entorno de desarrollo integrado (IDE) como Rstudio

Paquetes R

- Una gran cantidad de "**paquetes**" complementarios contribuido por los usuarios a lo largo de los años uno de los principales puntos fuertes de R.
- Muchos estadísticos han adoptado R como **su principal plataforma de programación**
- Muchos **Métodos estadísticos avanzados** disponibles en R
- Red integral de archivos R (CRAN): **repositorio para paquetes R**

Conceptos Importantes

- R ES UN LENGUAJE DE PROGRAMACIÓN DE ALTO NIVEL
- Rstudio es una interfaz
- Se basa en funciones
 - Función (argumentos)
 - Ejemplo: mean(datos\$X)

Paquetes:



[CRAN
Mirrors](#)
[What's new?](#)
[Task Views](#)
[Search](#)

[About R](#)
[R Homepage](#)
[The R Journal](#)

[Software](#)
[R Sources](#)
[R Binaries](#)
[Packages](#)
[Other](#)

[Documentation](#)
[Manuals](#)
[FAQs](#)
[Contributed](#)

Available CRAN Packages By Name	
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z	
A3	Accurate, Adaptable, and Accessible Error Metrics for Predictive Models
aaSEA	Amino Acid Substitution Effect Analyser
AATtools	Reliability and Scoring Routines for the Approach-Avoidance Task
ABACUS	Apps Based Activities for Communicating and Understanding Statistics
abbyyR	Access to Abbyy Optical Character Recognition (OCR) API
abc	Tools for Approximate Bayesian Computation (ABC)
abc.data	Data Only: Tools for Approximate Bayesian Computation (ABC)
ABC.RAP	Array Based CpG Region Analysis Pipeline
abcADM	Fit Accumulated Damage Models and Estimate Reliability using ABC
ABCanalysis	Computed ABC Analysis
abcdeFBA	ABCDE_FBA: A-Biologist-Can-Do-Everything of Flux Balance Analysis with this package
ABCOptim	Implementation of Artificial Bee Colony (ABC) Optimization
ABCp2	Approximate Bayesian Computational Model for Estimating P2
abcrf	Approximate Bayesian Computation via Random Forests
abcrlda	Asymptotically Bias-Corrected Regularized Linear Discriminant Analysis
abctools	Tools for ABC Analyses
abd	The Analysis of Biological Data
abdiv	Alpha and Beta Diversity Measures
abe	Augmented Backward Elimination
abess	Adaptive Best Subset Selection in Polynomial Time
abf2	Load Gap-Free Axon ABF2 Files
abglasso	Adaptive Bayesian Graphical Lasso
ABHgenotypeR	Easy Visualization of ABH Genotypes

Instalación de paquetes

- Usando código:
 - En la consola de R dentro de RStudio, utilizar el siguiente código:
`install.packages ("nombre del paquete")`
 - Ejecutar la linea
 - El proceso de instalación de paquetes puede tomar varios minutos.
- Usando ventana de gráficos
 - Seleccionar la pestaña paquetes(Packages) en el panel inferior izquierdo
 - Seleccionar Install
 - Escribir el nombre del paquete y dar instalar

Conceptos Importantes

- Se puede trabajar en la **consola** o en un “**script**”
- “Script” es un programa que puede guardarse
 - Se puede modificar fácilmente
 - Ayuda a la reproductibilidad de los resultados
 - Como todo programa puede ser “corrido” por línea o completo
- Los resultados pueden ser almacenado en un objeto y visualizarse en cualquier momento

Conceptos Importantes

- Una Fortaleza importante son los paquetes
 - Paquetes = rutinas generados por los usuarios
 - Ejemplos: `ggplot2` y `nmlm`
- Los paquetes se instalan una sola vez
 - Ejemplo: `install.packages ("ggplot2")`
- Los paquetes deben activarse cuando se van a usar
 - Solo una vez por sesión
 - ejemplo: `library(ggplot2)`

Sintaxis de R

- R es un lenguaje orientado a objetos.
- Los objetos son el elemento base del R.
- Los objetos se guardan con un nombre específico que no debe tener espacios.
- Un número, una matriz, en la gráfica, una función o una base de datos son ejemplos de objetos de R.
- El operador <- (=) sirve para asignar un objeto.

grafica <- plot (datos)

saludo = "Hola a todos" #cadenas de texto entre comillas

Practica 1:

- Revisar instalación de R-UCA
- Revisar instalación de R y Rstudio
- Crear proyecto y directorio de trabajo

Elementos básicos

R como calculadora

```
>3+2  
[1] 5  
> sqrt(10)  
[1] 3.162278  
> 4^2  
[1] 16  
> pi  
[1] 3.141593  
> sin(pi/2)  
[1] 1
```

Tipos de Archivo “nativos” de R

Datos	(.rdata)
Scripts	(.r)

• Primeros Comandos

```
ls()  
rm()  
save()  
load()
```

• Generar un objeto

Para generar un objeto, utilizamos “`<-`”
Opcionalmente se puede usar “`->`” y “`=`”
en la mayoría de la literatura no se usan

```
> a<-1  
> b<-2  
> d<-4  
> a+b  
[1] 3  
> (a+b)/d  
[1] 0.75  
> a+b/d  
[1] 1.5
```

Tipos de datos

- Cuando hablamos de tipos de datos en R, nos referimos a los objetos de datos más simples que podemos manejar, se conocen también como tipos de datos atómicos R
- En R hay 5 tipos de datos atómicos:
 - Carácter (character)
 - Reales (double)
 - Enteros (integer)
 - Complejos (complex)
 - Booleanos (logical)

Tipo de objetos en R

- **Vectores:** Variable unidimensional, todos los valores del mismo tipo;
- **Matrices:** Vector organizado en filas y columnas;
- **Arrays:** Es una generalización mutidimensional del vector
• (elementos del mismo tipo)
- **Listas:** lista de objetos que pueden ser de diferentes tipos y longitudes
- **Data frames:** conjunto de datos organizado en columnas de la misma longitud, pero puede tener diferente tipo y filas

Practica 1: R como calculadora(Tipo de variables)

- Construir un proyecto para el curso:
- Se recomienda primero fijar el directorio de trabajo:
 - `setwd("C:/Users/Gustavo/Desktop/output R/C_Calidad")`
- Abrir un nuevo archivo de tipo “script”
- Trabajar en el programa, al finalizar guardar el script

Tabla símbolos básicos

Aritméticos	Lógicos	Comparativos	
+	Adición	==	Igual a \& Y Lógico
-	Substracción	!=	Diferente de !
*	Multiplicación	<	Menor que O Lógico
/	División	>	Mayor que is.na(x) ¿Es NA?
^	Potencia	<=	Menor o igual a is.null(x) ¿Es Null?
\%/\%	División de enteros	>=	Mayor o igual a is.nan(x) ¿Es NaN?

Tabla de funciones de distribución

Distribución	Función de densidad	distribución acumulada	función cuantil	Uso de números aleatorios
Normal	<code>dnorm(...)</code>	<code>pnorm(...)</code>	<code>qnorm(...)</code>	<code>rnorm(...)</code>
Uniforme	<code>dunif(...)</code>	<code>punif(...)</code>	<code>qunif(...)</code>	<code>runif(...)</code>
Student's t	<code>dt(...)</code>	<code>pt(...)</code>	<code>qt(...)</code>	<code>rt(...)</code>
Chi-cuadrada	<code>dchisq(...)</code>	<code>pchisq(...)</code>	<code>qchisq(...)</code>	<code>rchisq(...)</code>
Gamma	<code>dgamma(...)</code>	<code>pgamma(...)</code>	<code>qgamma(...)</code>	<code>rgamma(...)</code>
Normal Multivariada	<code>dmvnorm(...)</code>	<code>pmvnorm(...)</code>		<code>rmvnorm(...)</code>
Student's t Multivariada	<code>dmvt(...)</code>	<code>pmvt(...)</code>	<code>qmvt(...)</code>	<code>rmvt(...)</code>
Exponencial	<code>dexp(...)</code>	<code>pexp(...)</code>	<code>qexp(...)</code>	<code>rexp(...)</code>
Cauchy	<code>dcauchy(...)</code>	<code>pcauchy(...)</code>	<code>qcauchy(...)</code>	<code>rcauchy(...)</code>
Wishart	<code>dwish(...)</code>			<code>rwish(...)</code>

Tabla funciones básicas

Función	Matemáticas	Función	Estadísticas
sqrt(x)	Raíz de x	mean(x)	Media
exp(x)	Exponencial de x	sd(x)	Desviación estándar
log(x)	Logaritmo natural de x	var(x)	Varianza
log10(x)	Logaritmo base 10	median(x)	Mediana
sum(x)	Suma de los elementos de x	quantiles(x)	Quantiles
prod(x)	Producto de los elementos de x	cor(x,y)	Correlación
sin(x)	Seno	max(x)	Valor máximo
cos(x)	Coseno	min(x)	Valor mínimo
tan(x)	Tangente	range(x)	Retorna el máximo y mínimo
round(x,n)	Redondea a n dígitos	sort(x)	Ordena los elementos de x
cumsum(x)	Calcula las sumas acumuladas	summary	Resumen de las variables

Practica 2: Tipo de objetos

- En esta práctica se identifican los tipos de objetos (vectores, matrices, listas y data frames).
- Se muestra cómo puede construirse, manipularse y realizar operaciones con los distintos tipos de objetos

Datos: U. S. State Public-School Expenditures

- Description: The Anscombe data frame has 51 rows and 4 columns. The observations are the U. S. states plus Washington, D. C. in 1970.
- This data frame contains the following columns:
- **Education:** Per-capita education expenditures, dollars.
- **Income:** Per-capita income, dollars.
- **Young:** Proportion under 18, per 1000.
- **Urban:** Proportion urban, per 1000.
- **Source:** Anscombe, F. J. (1981) Computing in Statistical Science Through APL. Springer-Verlag.