

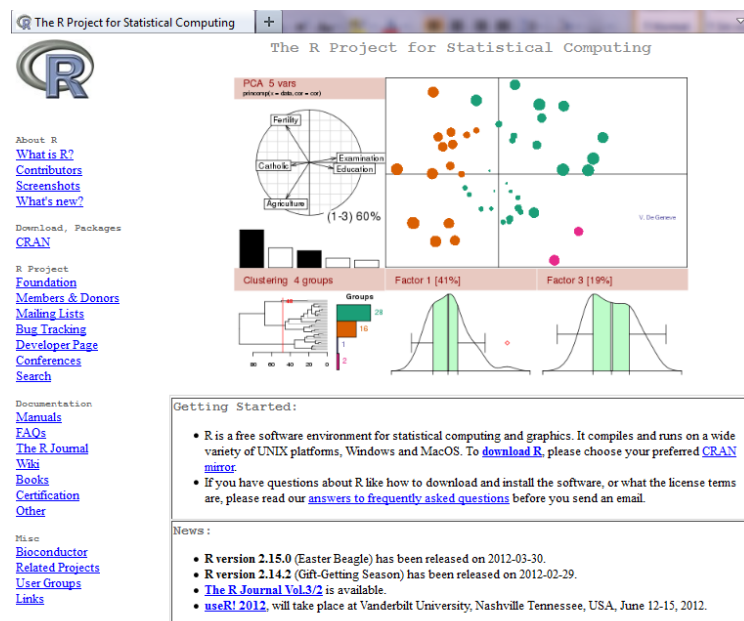


## PRÁCTICA 1. INTRODUCCIÓN A R-COMANDER. INSTALACIÓN DEL PAQUETE Rcmdr. INTRODUCCIÓN DE DATOS, OPERACIONES CON VARIABLES Y PRIMEROS GRÁFICOS PARA VARIABLES CUALITATIVAS.

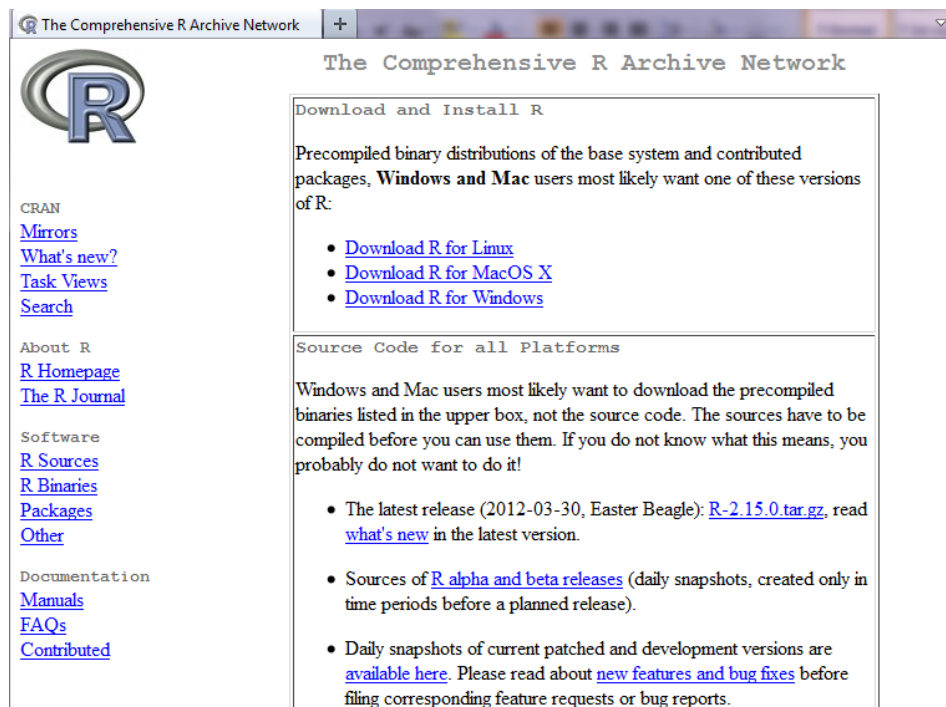
### 1.- INSTALACIÓN

**R** es un lenguaje de programación y un entorno para análisis estadístico y la realización de gráficos. Fue inicialmente escrito por Robert Gentleman y Ross Ihaka del Departamento de Estadística de la Universidad de Auckland en Nueva Zelanda. **R** es un proyecto GNU, similar a su versión comercial, S, desarrollada por los laboratorios Bell. Desde 1997, se formó lo que se conoce como núcleo de desarrollo de **R**, que es quien tiene actualmente la posibilidad de modificación directa del código fuente de **R**. No obstante, **R** es actualmente resultado de un esfuerzo de colaboración de personas en todo el mundo. El código de **R** está disponible como software libre bajo las condiciones de licencia GNU-GPL. La página web del proyecto es <http://www.r-project.org>. En el presente curso, nos centramos por su sencillez y facilidad de empleo, en la interfaz gráfica de **R**, conocida como **R-Commander**.

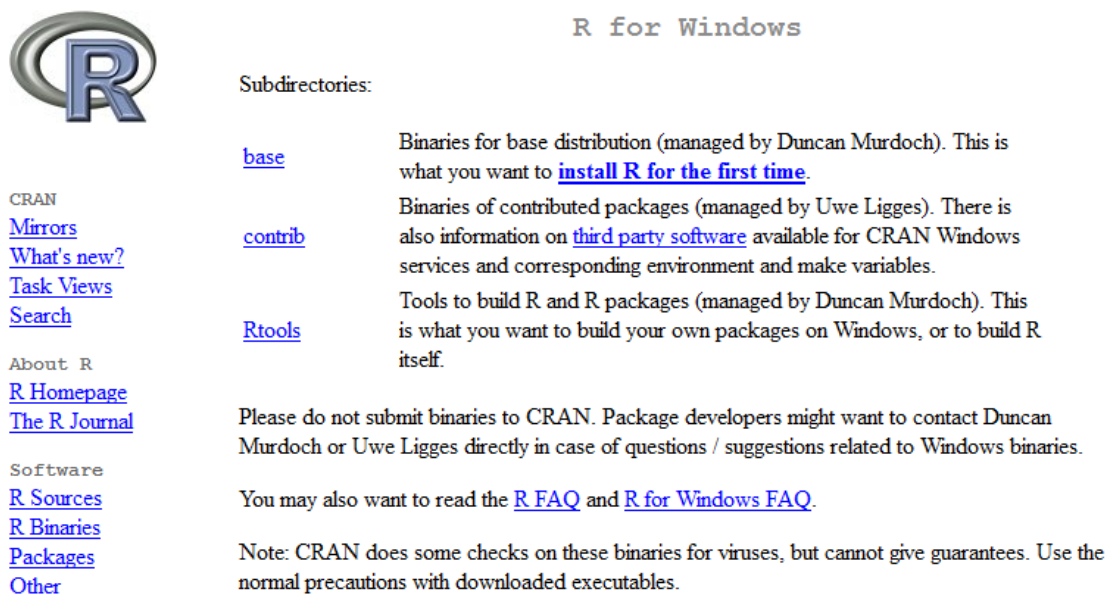
#### Instalación de R



Se selecciona el país desde donde se desea descargar el programa pulsando **CRAN mirror**. Una vez seleccionado España se elige el sistema operativo, como se muestra a continuación



Se selecciona **base** para acceder a la ventana de descarga del programa



donde se puede descargar la versión 2.15.0 de R disponible.





## R-2.15.0 for Windows (32/64 bit)

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](#)

[Download R 2.15.0 for Windows](#) (47 megabytes, 32/64 bit)

[Installation and other instructions](#)

New features in this version: [Windows specific](#), [all platforms](#).

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

### Frequently asked questions

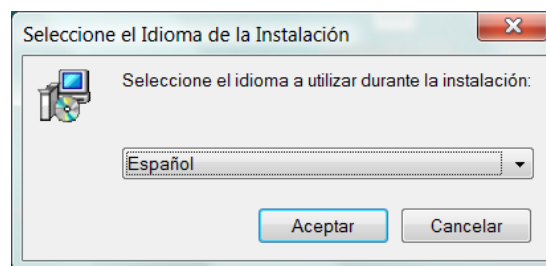
- [How do I install R when using Windows Vista?](#)
- [How do I update packages in my previous version of R?](#)
- [Should I run 32-bit or 64-bit R?](#)

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

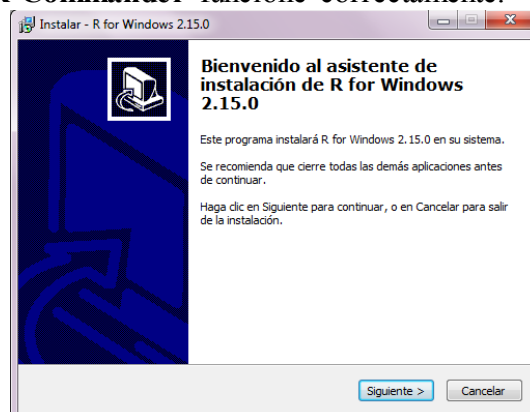
### Other builds

- 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](#).

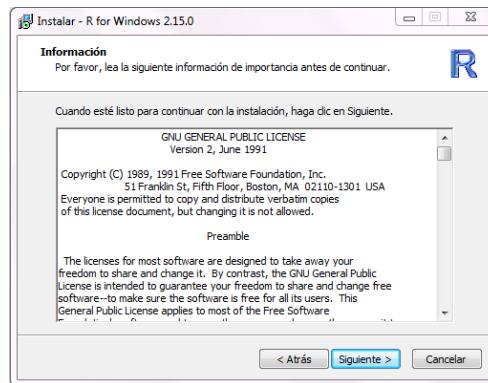
Una vez descargado el fichero ejecutable de **R**, se instala en el ordenador. El primer paso es la selección del idioma que se va a emplear en la instalación. El idioma en el que aparecen las ventanas y menús de **R** y **R-Commander** es seleccionado de forma automática del idioma definido en el sistema operativo.



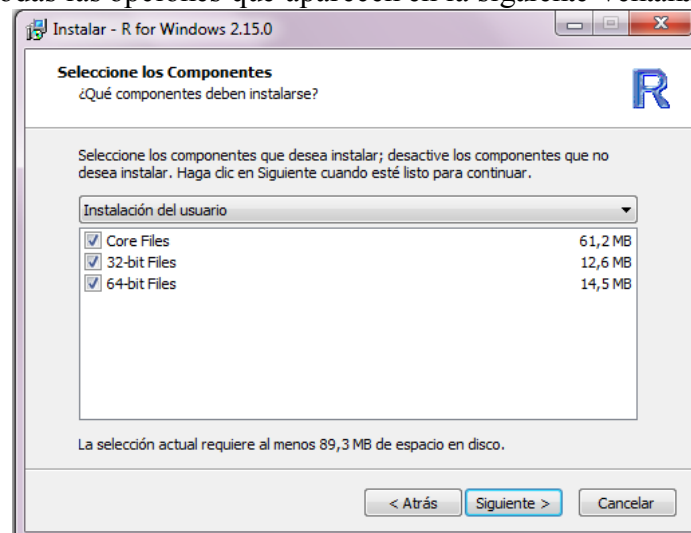
Se especifican a continuación algunas opciones de instalación del programa, que son importantes para que **R-Commander** funcione correctamente.



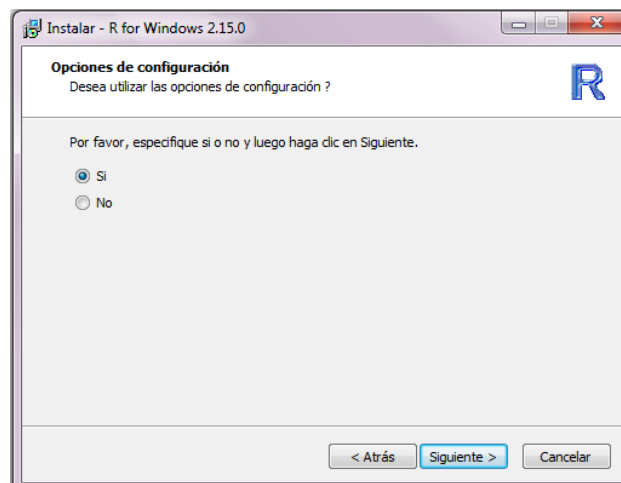
Se pulsa **Siguiente**

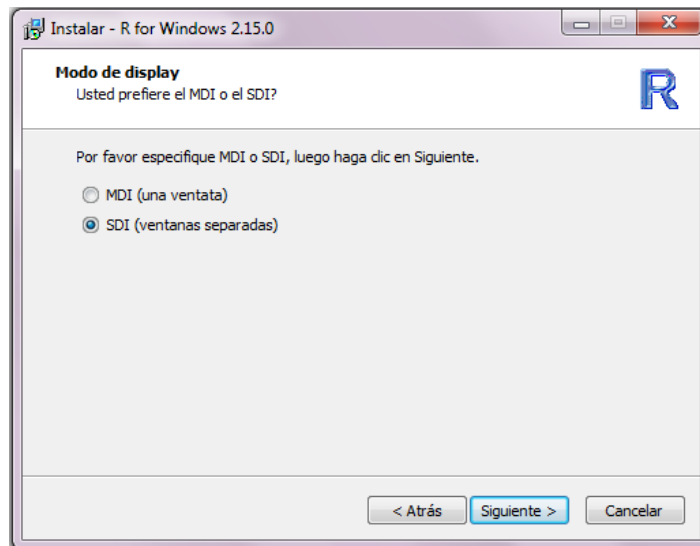


Seleccionamos todas las opciones que aparecen en la siguiente ventana:

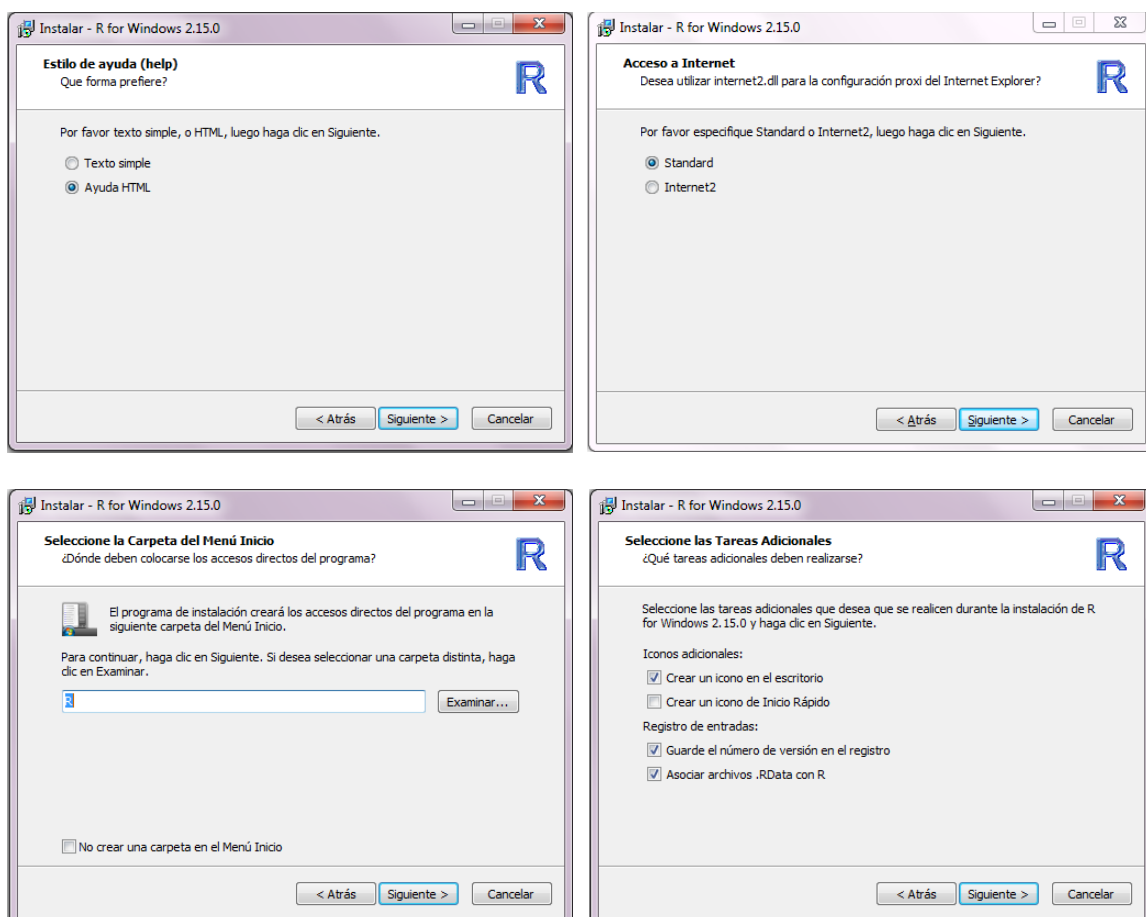


Para que **R-Commander** funcione correctamente es conveniente seleccionar la opción **Sí** en la ventana de opciones de configuración, así como **SDI, ventanas separadas**.





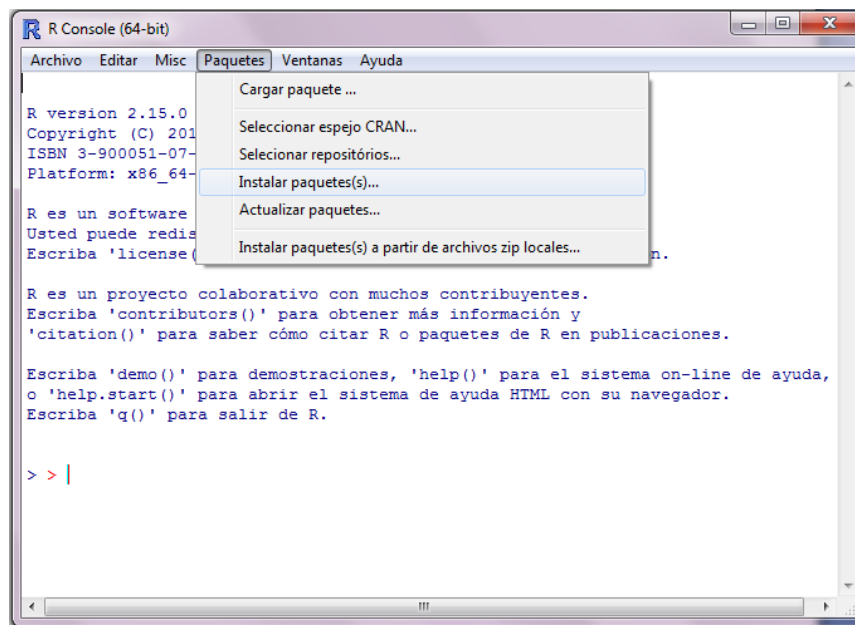
Se muestran pantallas con algunas opciones de instalación:



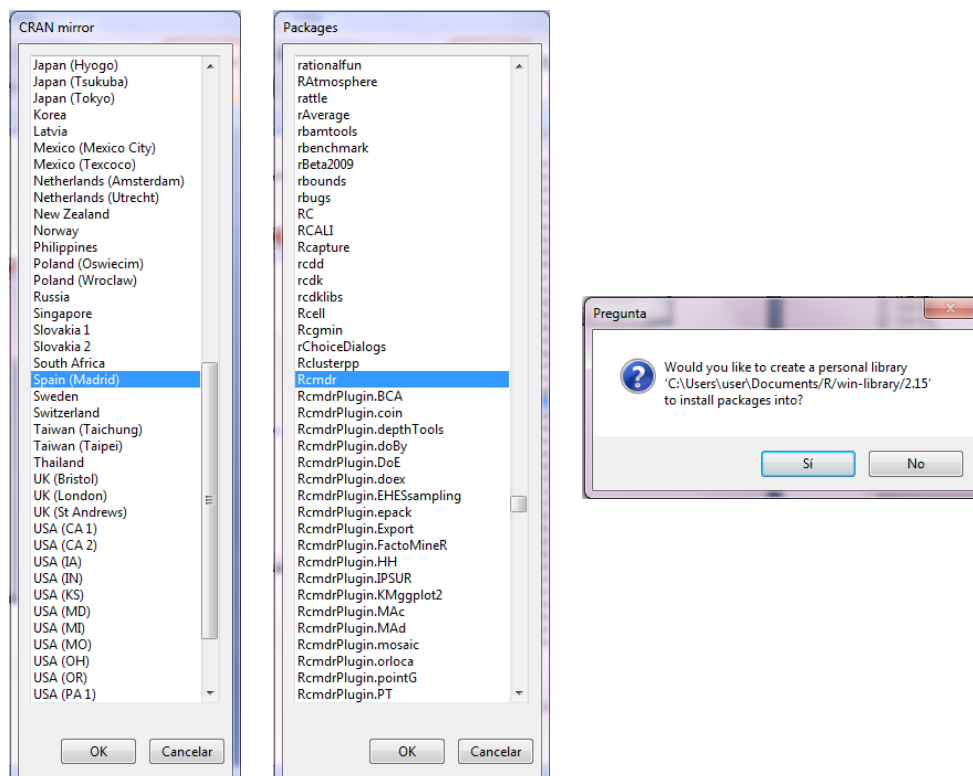
## Instalación de R-Commander



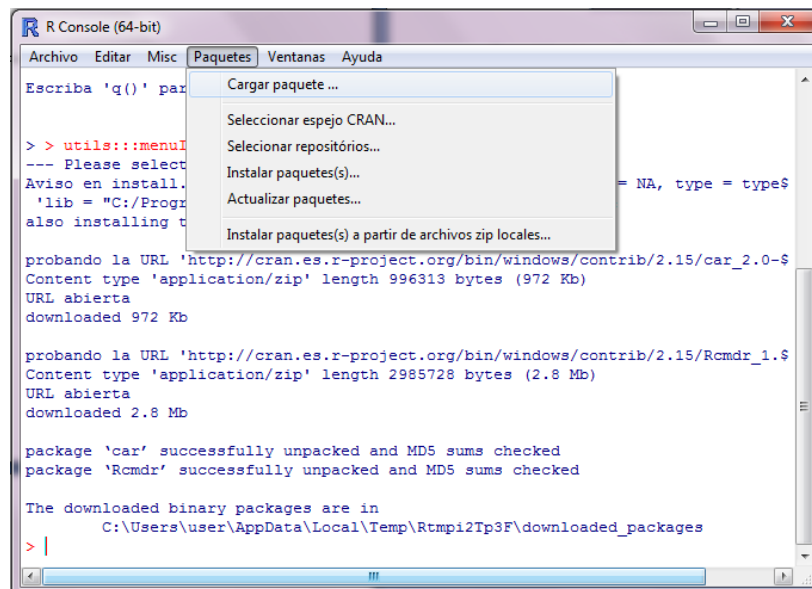
Una vez instalado **R**, se instala **R-Commander** seleccionando en el menú de **R** la librería de **R-Commander** como se muestra a continuación:



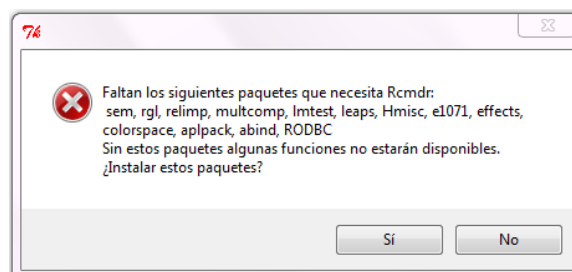
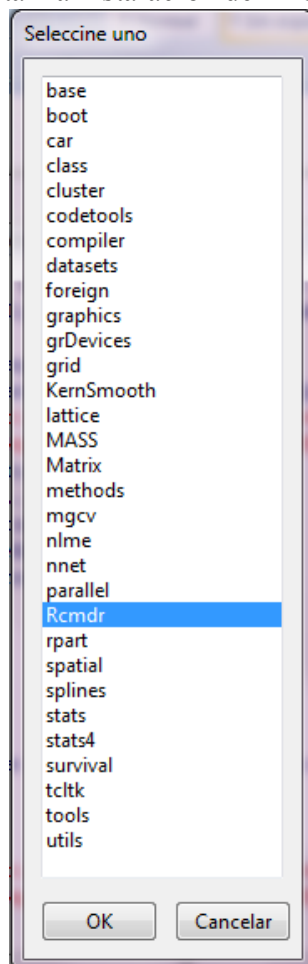
Se descarga **R-Commander** seleccionando el paquete **Rcmdr**. Es necesario crear una librería personal con los paquetes que instalemos

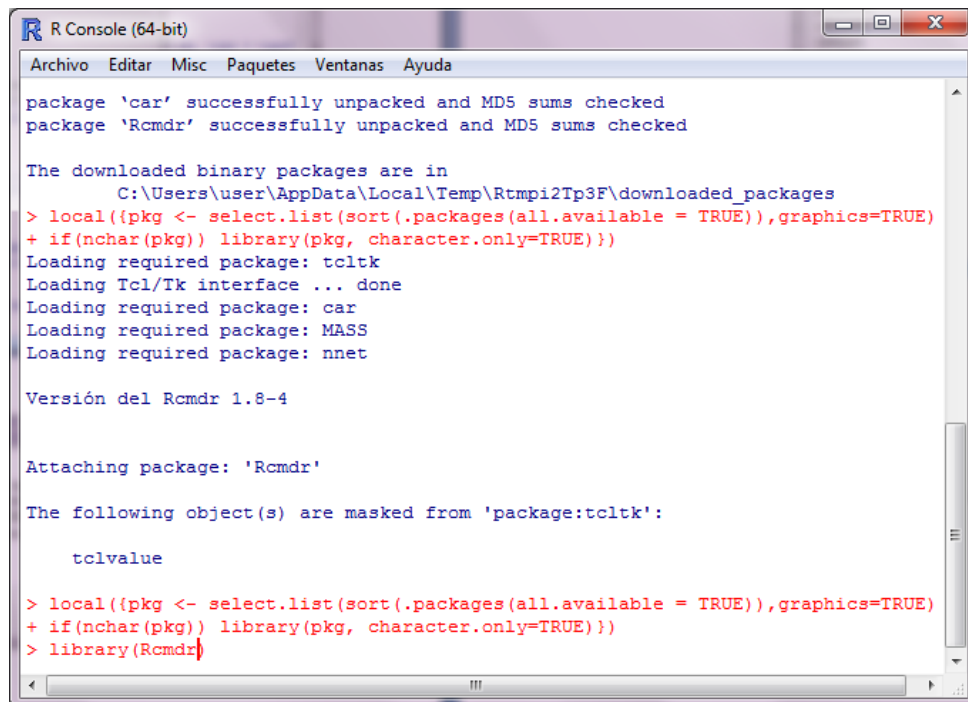


Una vez instalado el paquete se carga en el menú de **R** como se muestra a continuación:



Se selecciona de nuevo el paquete **Rcmdr** y se instalan los paquetes que faltan para completar la instalación de **R-Commander**





```
R Console (64-bit)
Archivo  Editar  Misc  Paquetes  Ventanas  Ayuda

package 'car' successfully unpacked and MD5 sums checked
package 'Rcmdr' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:\Users\user\AppData\Local\Temp\Rtmpi2Tp3F\downloaded_packages
> local({pkg <- select.list(sort(.packages(all.available = TRUE)),graphics=TRUE)
+ if(nchar(pkg)) library(pkg, character.only=TRUE)})
Loading required package: tcltk
Loading Tcl/Tk interface ... done
Loading required package: car
Loading required package: MASS
Loading required package: nnet

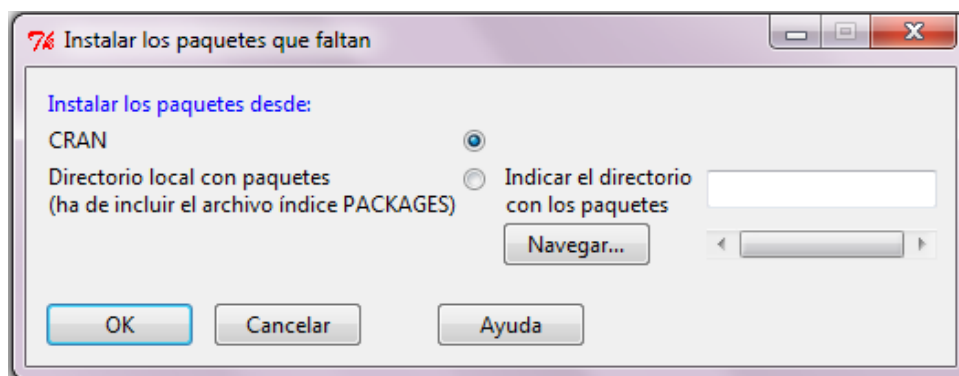
Versión del Rcmdr 1.8-4

Attaching package: 'Rcmdr'

The following object(s) are masked from 'package:tcltk':

  tclvalue

> local({pkg <- select.list(sort(.packages(all.available = TRUE)),graphics=TRUE)
+ if(nchar(pkg)) library(pkg, character.only=TRUE)})
> library(Rcmdr)
```



Con esta instalación final de paquetes desde **CRAN** finaliza la instalación de **R-Commander**.

Para iniciar **R-Commander** es necesario escribir la orden **library(Rcmdr)** al inicio de la consola de **R**. Si se cierra **R-commander**, pero no **R**, basta con escribir **Commander()** para abrir **R-commander** de nuevo.

En la siguiente figura se muestra el aspecto general de **R-Commander**:







Las ventanas **R-Commander** y de la consola **R** flotan en el escritorio. Los menús y cuadros de diálogo de **R-Commander** se usan para leer, manipular y analizar datos. Las instrucciones **R** generadas por la GUI de **R-Commander** aparecen en la ventana de texto superior (denominada Ventana de Instrucciones) dentro de la ventana principal de **R-Commander**. El resultado que proporciona **R-commander** aparece en la segunda ventana de texto (denominada Ventana de Resultados). La ventana gris inferior (denominada Mensajes) muestra mensajes de error, advertencias y otra información. Cuando se crea un gráfico aparece en una ventana separada llamada Graphics Device.



## 2.- INTRODUCCIÓN DE DATOS. PRIMEROS GRÁFICOS

### Ejercicio 1

Ejecute library(Rcmdr). Trabajaremos con el menú “Datos”

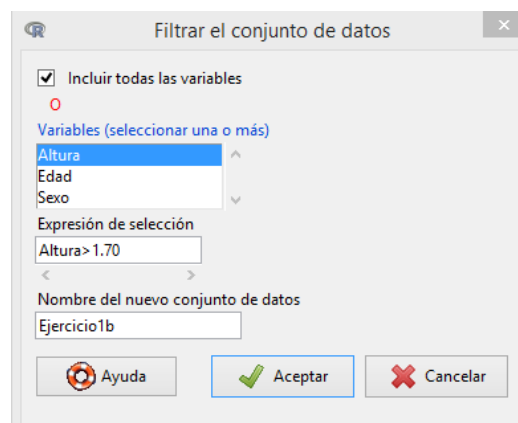
Se han medido las alturas de un grupo de alumnos.

1.75	1.80	1.77	1.75	1.71	1.64	1.69	1.95	1.75	1.67
------	------	------	------	------	------	------	------	------	------

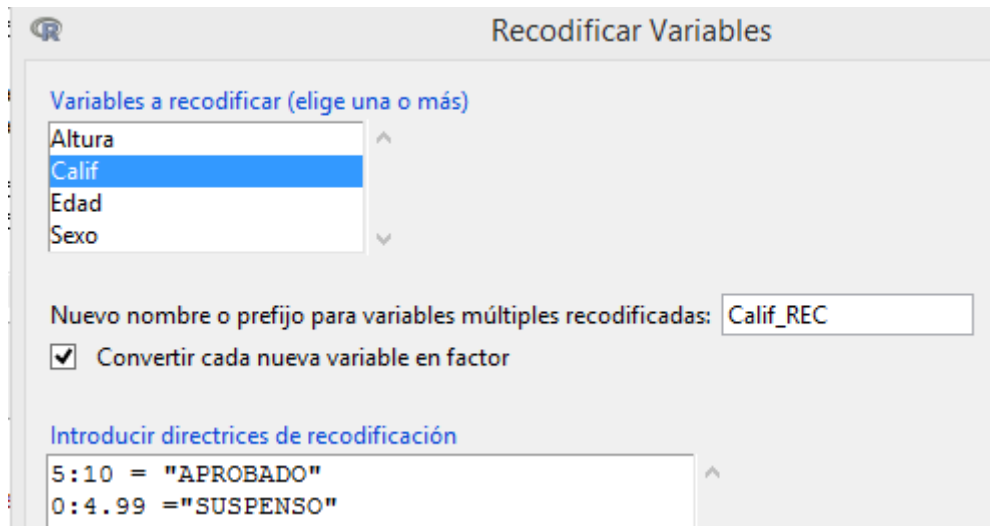
- Cree una carpeta en el escritorio llamada “Practical” (Menú Fichero). Introducimos los datos de esta variable denominada “Altura” en un fichero denominado “Ejercicio1”. Cambiamos el nombre a la variable.
- Guardamos el conjunto de datos activo (Menú Datos/Conjunto de datos activo) en la carpeta “Practical”.
- Añadimos otras 4 variables (Nombre, Edad, Sexo, Calificación). No le ponga acentos a los nombres de las variables. Puede llamar a la última Calif:

Juan	Salvador	Luis	José	Lola	Daniel	David	Carlos	Agustín	Maite
20	21	20	19	19	18	20	18	19	19
H	H	H	H	M	H	H	H	H	M
6.3	9.1	5.2	3.1	5.8	8.6	2.2	6.3	9.3	9.5

- Guardamos de nuevo el fichero y miramos el nombre de las variables (Menú Datos/Conjunto de datos activo/Variables...). Nos aparecen los nombres de las variables.
- Establecemos los nombres de los casos por la variable “Nombre” (Menú Datos/Conjunto de datos activo/Establecer nombres de casos). Deben ser todos distintos para poder realizar esta operación.
- Filtramos el conjunto de datos para obtener un fichero que contenga a los alumnos con altura superior a 1.70. Visualice dicho fichero y guárdelo “Ejercicio1b”.

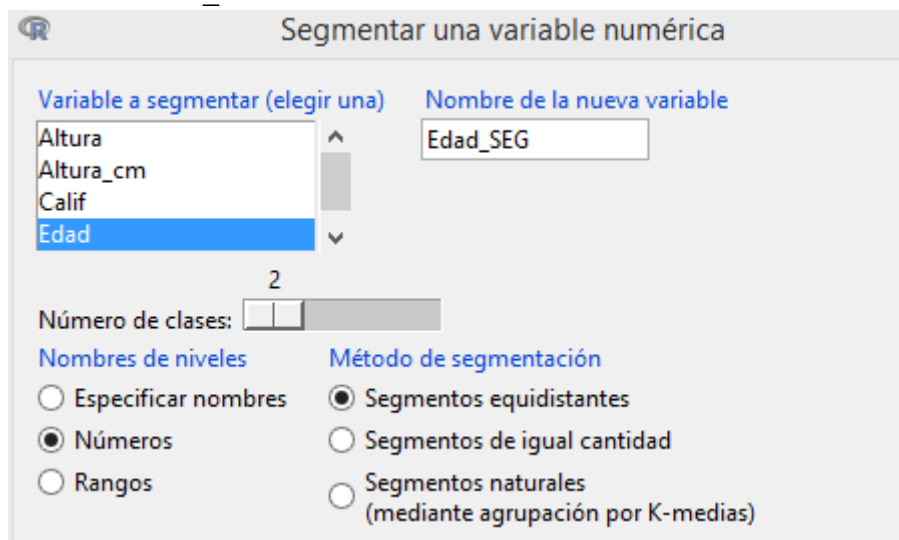


- g) Seleccione como conjunto de datos activo el “Ejercicio1” (Menú Datos/Conjunto de datos activo/Seleccionar conjunto de datos activo).
- h) Vamos ahora a recodificar la variable “Calificación” para conocer los alumnos SUSPENSOS y APROBADOS (Menú Datos/Modificar variables/Recodificar).



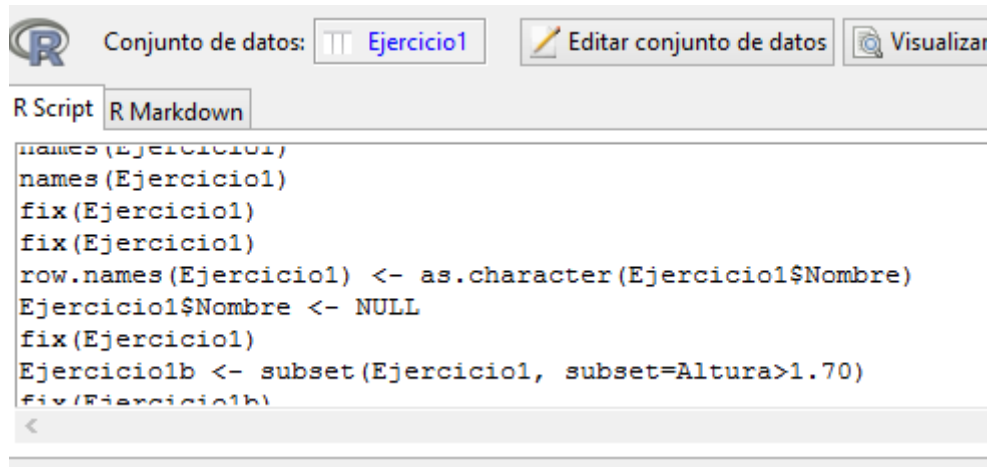
Visualice el conjunto de Datos para comprobar que ha realizado correctamente esta operación.

- i) Supongamos ahora que necesitamos la Altura en cm. Vamos a calcular una nueva variable Altura\_cm, multiplicando la original por 100 (Menú Datos/Modificar variables/Calcular).
- j) Ejecute (Menú Datos/Modificar variables/Añadir número de observaciones). Compruebe que ha añadido la columna “ObsNumber”.
- k) Vamos a segmentar la variable “Edad”, para formar automáticamente grupos de observaciones, que más adelante podremos utilizar para realizar por ejemplo, gráficos por subgrupos de casos (Menú Datos/Modificar variables/Segmentar). En este caso, formaremos solamente dos grupos y llamaremos a la nueva variable “Edad\_SEG”.



1) Guarde el fichero de datos.

Nota: Observe que todas las órdenes que hemos ejecutado aparecen en el Script. Puede modificar alguna de ellas y ejecutarla sin tener que acceder al Menú. Este entorno es más parecido al de R, en el que solamente se trabaja por comandos.

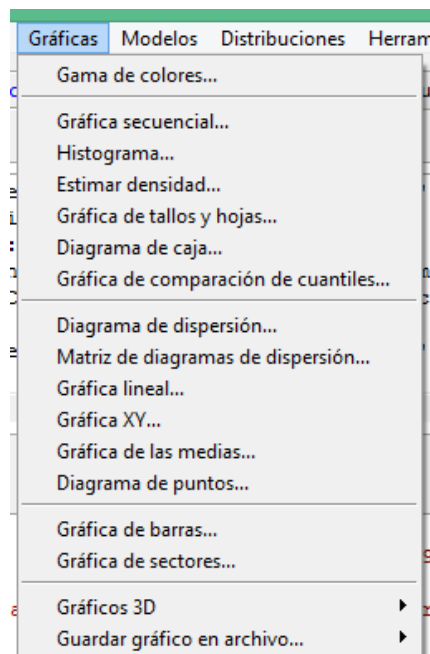


```
names(Ejercicio1)
names(Ejercicio1)
fix(Ejercicio1)
fix(Ejercicio1)
row.names(Ejercicio1) <- as.character(Ejercicio1$Nombre)
Ejercicio1$Nombre <- NULL
fix(Ejercicio1)
Ejercicio1b <- subset(Ejercicio1, subset=Altura>1.70)
fix(Ejercicio1b)
```

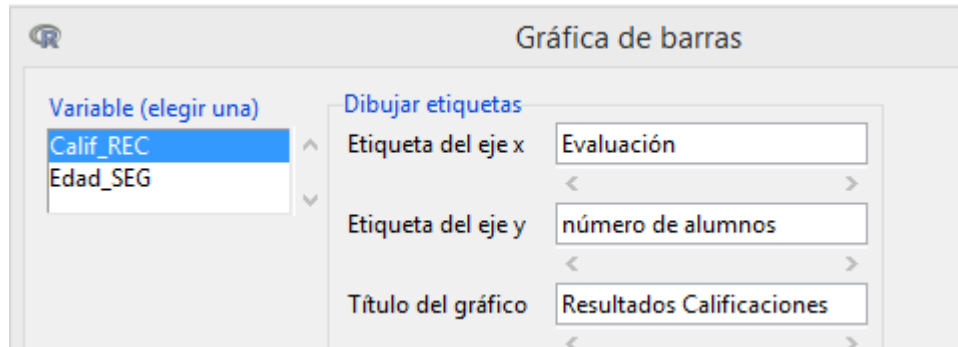
## Ejercicio 2

Ejecute `library(Rcmdr)`. Trabajaremos con el menú “Gráficos”

Utilizaremos los datos del Ejercicio 1 para realizar algunos gráficos de las variables que hemos creado. El menú “Gráficas” contiene las siguientes opciones:



- a) En el menú datos cargue el conjunto de datos “Ejercicio1.RData”. Visualice su contenido. Debe tener las siguientes columnas: "Altura", "Edad", "Sexo", "Calif", "Nombre", "Calif\_REC", "Altura\_cm", "ObsNumber", "Edad\_SEG".
- b) Cree un diagrama de barras con la columna “Calif\_REC”, como sigue:



Nos crea un gráfico de barras con los resultados de las calificaciones de los alumnos. En el Script nos aparece:

```
barplot(table(Ejercicio1$Calif_REC), xlab="Evaluación", ylab="número de alumnos", main="Resultados Calificaciones").
```

Este comando tiene múltiples opciones de mejora. Para mayor información:

<https://stat.ethz.ch/R-manual/R-devel/library/graphics/html/barplot.html>

Si queremos dar color al gráfico (p.e. verde), hemos de introducir en la orden anterior:

```
barplot(table(Ejercicio1$Calif_REC), xlab="Evaluación", ylab="número de alumnos", main="Resultados Calificaciones", col="green").
```

Puede encontrar la gama de colores disponibles en:

<http://www.stat.columbia.edu/~tzheng/files/Rcolor.pdf>

- c) Realice pruebas con otros colores. También puede cambiar los colores del título, ejes, etiquetas del eje, etc. Pruebe lo siguiente e identifique cada opción:

```
barplot(table(Ejercicio1$Calif_REC), xlab="Evaluación", ylab="número de alumnos", main="Resultados Calificaciones", col = "blueviolet", border = "dark blue", horiz=TRUE, col.main="coral4", col.lab="deeppink4", col.axis="green4")
```

- d) R también tiene algunas gamas de colores predefinidos. Puede verlos en:

<http://stat.ethz.ch/R-manual/R-devel/library/grDevices/html/palettes.html>

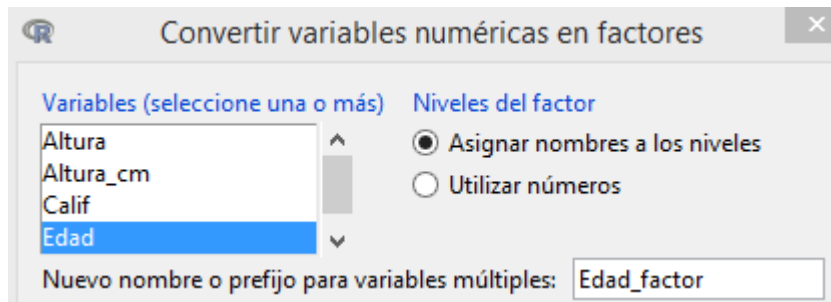
En este caso, pruebe con la orden:

```
barplot(table(Ejercicio1$Calif_REC), xlab="Evaluación", ylab="número de alumnos", main="Resultados Calificaciones", col=heat.colors(5))
```

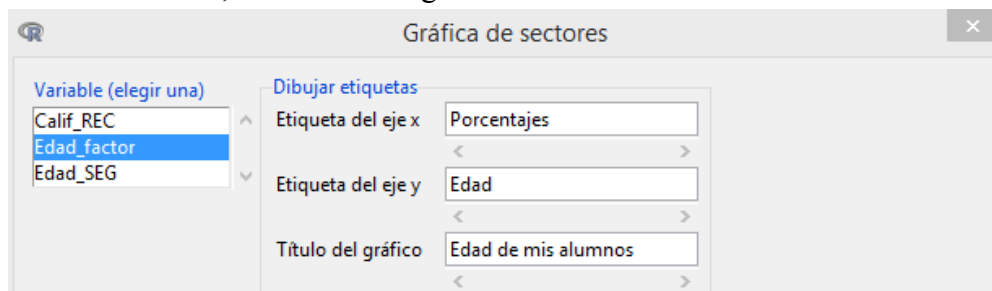
- e) Realice ahora un gráfico de sectores y pruebe las distintas opciones que tiene.



- f) Vamos a realizar ahora un gráfico de sectores para la variable “Edad” en el que vengan expresados los porcentajes observados. En primer lugar, hemos de convertir dicha variable en factor (cualitativa). En el Menú Datos podemos convertir la variable numérica en factor.



Y a continuación, realizar el diagrama de sectores:



Pongamos ahora las etiquetas (con porcentajes) a cada uno de los sectores. Escribimos en el Script lo siguiente para crear una tabla de frecuencias para la edad:

```
tabla<-table(Ejercicio1$Edad_factor)
```

Calculemos los porcentajes de la tabla y redondeo a un solo decimal:

```
porcentajes<-round(prop.table(tabla)*100,1)
```

Incluyamos las etiquetas:

```
etiquetas<-c("18","19","20","21")
```

```
etiquetas<-paste(etiquetas, "años, ", porcentajes, "%")
```

```
etiquetas
```

Realicemos el diagrama de sectores, incluyendo dichas etiquetas:

```
pie(table(Ejercicio1$Edad_factor), labels=etiquetas, xlab="Porcentajes",  
ylab="Edad", main="Edad de mis alumnos",  
col=rainbow_hcl(length(levels(Ejercicio1$Edad_factor))))
```

