



CEU

*Universidad  
Cardenal Herrera*

***Instalación e introducción al  
software estadístico R y la librería  
R-Commander.  
Estadística descriptiva.***

Botella-Rocamora, P.; Alacreu-García, M.; Martínez-Beneito, M.A.

# ÍNDICE

	<b>Página</b>
1. Instalación e introducción al software estadístico R.	<b>3</b>
a. Introducción .....	3
b. Instalación del programa base R .....	4
c. Instalación y acceso a R-Commander .....	6
d. Funciones básicas de R-Commander .....	8
2. Estadística descriptiva con R-Commander .....	<b>16</b>
a. Introducción .....	16
b. Resumen numérico de variables cuantitativas y categóricas .....	17
c. Resumen gráfico de variables cuantitativas y categóricas .....	20
d. Resumen numérico/gráfico bivalente	27



CEU

Universidad  
Cardenal Herrera

# 1. Instalación e introducción al software estadístico R.

## *a. Introducción*

R es un software libre que permite realizar análisis estadísticos y el más usado en la comunidad científica. Este programa está disponible en la página web: <http://www.r-project.org> y consta de una aplicación central y de librerías de multitud de temas que se pueden instalar según necesidad.

R es un programa de instrucciones, y por tanto, no resulta del todo "amigable" para los usuarios que no están acostumbrados a este tipo de manejo. Actualmente existe una interfaz que permite el manejo del programa R mediante una ventana de menús, este interfaz se llama **R-Commander**. Este interfaz permite al usuario comenzar a manejar este programa sin conocer el lenguaje de instrucciones, y permite el aprendizaje de este lenguaje de forma sencilla (si el usuario así lo quiere).

En el libro ***Estadística Básica con R y R-Commander*** se puede aprender de forma sencilla, desde la instalación del programa hasta el manejo en análisis estadísticos sencillos. Este libro también es libre y se puede obtener en la página web [knuth.uca.es/ebrcmdr](http://knuth.uca.es/ebrcmdr).

En esta sección se desarrollarán los siguientes puntos:

- Instalación del programa y el interfaz R-Commander
  - Instalación del programa base R
  - Instalar librería "Rcmdr"
  - Acceso a la aplicación R-Commander
- Conceptos básicos
  - Directorio de trabajo
  - Ficheros de instrucciones
- Entorno de Datos
  - Crear fichero de datos
  - Importar fichero de datos (Excel y Access)
  - Manejo del fichero de datos
    - Modificación de variables del conjunto de datos.
    - Recodificación de una variable existente
    - Cálculo de una nueva variable
  - Cargar/Guardar ficheros de instrucciones



CEU

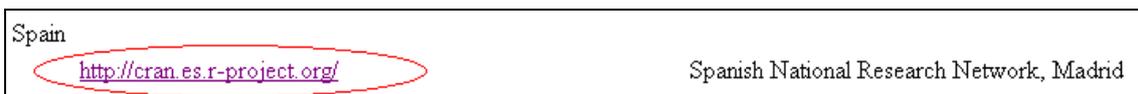
Universidad  
Cardenal Herrera

## ***b. Instalación del programa base R.***

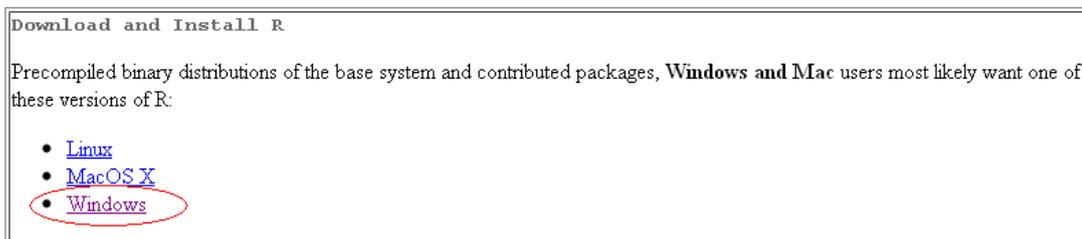
En la página web <http://www.r-project.org>, seleccionamos en la parte izquierda la opción Download, Packages pinchando en CRAN



Accedemos entonces a una ventana en la que nos aparecen diferentes lugares desde los que descargar el programa. Podemos seleccionar, por ejemplo, la página de nuestro país (aunque si no funciona del todo bien o está sobrecargada podemos acceder a cualquier otro lugar para descargar el programa)



Seleccionamos a continuación nuestro sistema operativo (en la Universidad: Windows)



Y a continuación seleccionamos el programa base con la opción:



R for Windows

This directory contains binaries for a base distribution and packages to run on i386/x64 Windows.

Note: CRAN does not have Windows systems and cannot check these binaries for viruses. Use the normal precautions with downloaded executables.

Subdirectories:

<a href="#">base</a>	Binaries for base distribution (managed by Duncan Murdoch)
<a href="#">contrib</a>	Binaries of contributed packages (managed by Uwe Ligges)

Por último pinchamos sobre la opción "Download R ....." para que comience la descarga del programa:

R-3.0.2 for Windows (32/64 bit)

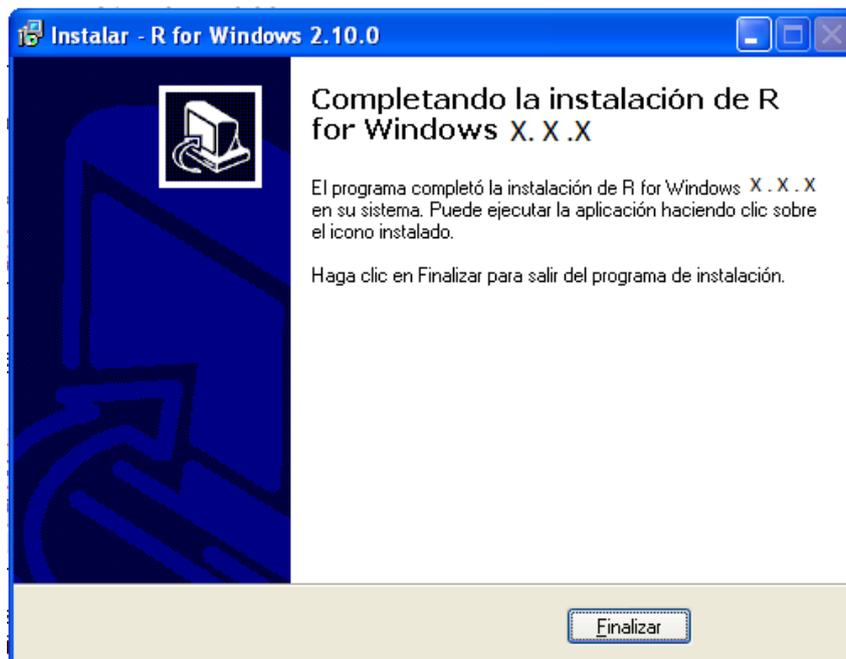
[Download R 3.0.2 for Windows](#) (52 megabytes, 32/64 bit)

[Installation and other instructions](#)

[New features in this version](#)

En ese momento se nos pide una carpeta en la que guardar el ejecutable que instalará la aplicación. Dependiendo del tamaño del fichero y de la conexión a internet este proceso puede tardar varios minutos.

Una vez descargado, ejecutamos el fichero descargado y comenzará la instalación del programa R. Durante el proceso se nos pedirá además del idioma para el proceso de instalación seleccionar algunas opciones. Se puede instalar pulsando siempre la opción "Siguiente" hasta llegar a "Finalizar":



El programa base ya está instalado.

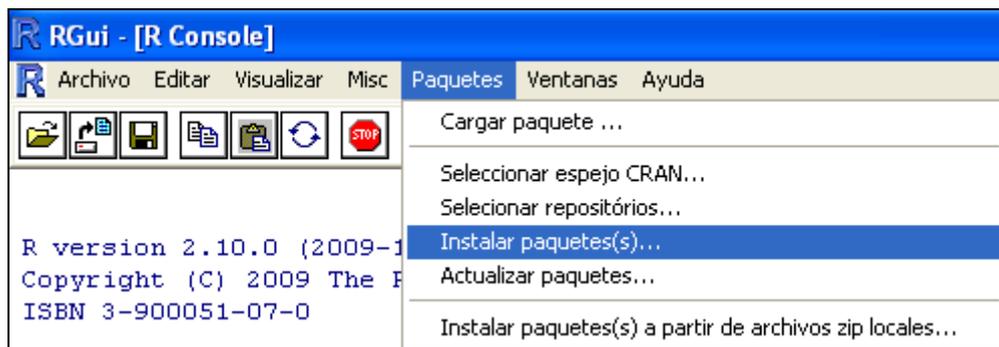
### **c. Instalación y acceso a la librería R-commander.**

“R commander” es una librería del programa R y para instalarla tenemos que arrancar en primer lugar el programa. Para ello, o bien buscamos el icono que por defecto se crea en el escritorio del ordenador, o bien buscamos en el menú de Windows> Programas>R X.XX.X



Al acceder al programa, aparecerá una ventana vacía en la que se pueden escribir instrucciones (comandos) y que tiene un menú de opciones en la parte superior.

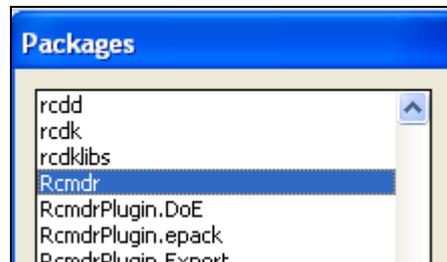
Para instalar la librería de “R commander” accedemos en el menú a:



A continuación seleccionamos de nuevo el lugar del que nos queremos descargar el paquete y pulsamos OK:



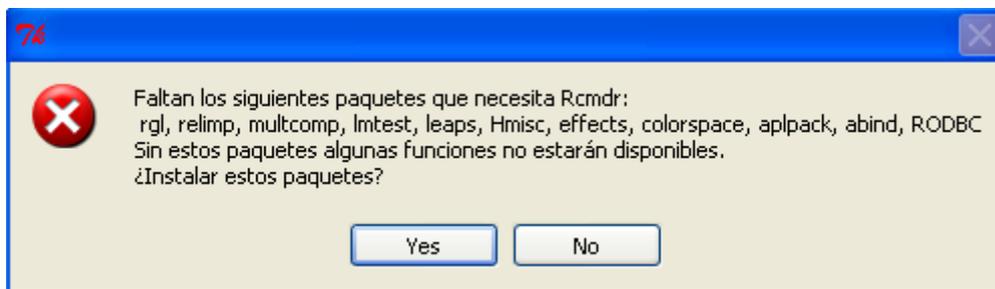
Y en la venta Packages que aparece buscaremos desplazándonos con la barra lateral hacia abajo el paquete **Rcmdr** y pulsamos OK:



Comenzará entonces la descarga e instalación de esta librería. Cuando finalice la descarga e instalación escribiremos al lado de la flecha roja ( > ) la sentencia `library(Rcmdr)`.

```
The downloaded packages are in
      C:\Documents and Settings\A
> library(Rcmdr)
```

Esta sentencia abrirá el interfaz del R commander. Pero si es la primera vez que lo intentamos cargar, nos dará el siguiente mensaje:



Nos avisa de otras librerías que necesita cargar para que el programa pueda funcionar. Respondemos "Yes" a este mensaje y nos pregunta de qué lugar puede obtener las librerías que le faltan: bien desde internet (opción CRAN que está seleccionada por defecto) o bien desde algún directorio local de nuestro ordenador (en el caso en el que nosotros nos hubiéramos descargado ya estas librerías con anterioridad desde internet). Dejamos la opción por defecto para que las descargue desde internet y pulsamos OK:



CEU

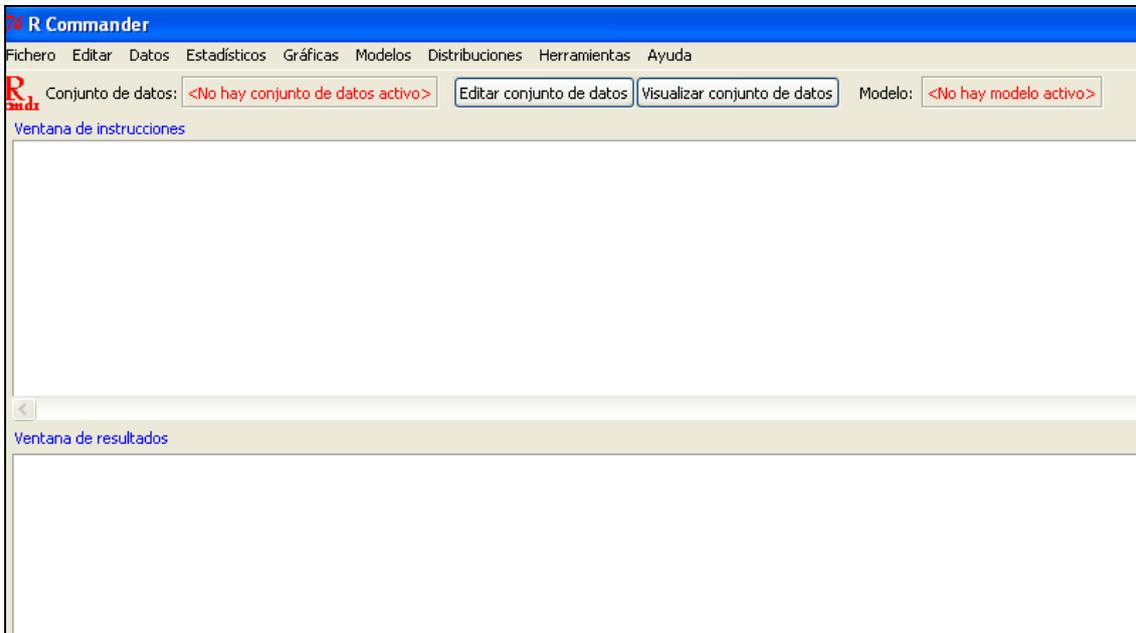
Universidad  
Cardenal Herrera



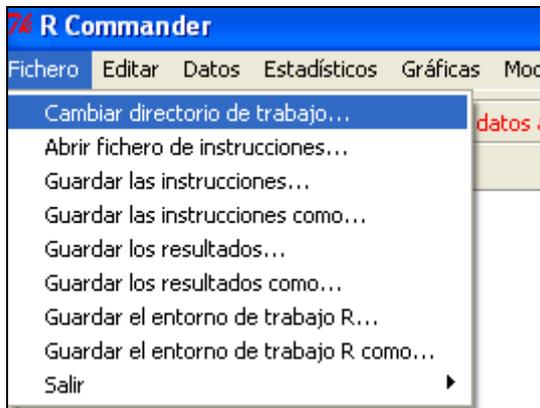
A partir de ese momento comienzan a descargarse e instalarse todas las librerías que el programa necesitará para funcionar (puede tardar varios minutos). Cuando finalice, se abrirá la ventana de **R commander**

Este proceso es necesario realizarlo la primera vez que se accede, pero en futuros accesos ya estará todo instalado y la entrada al programa será directa simplemente tecleando en la ventana que abre el programa R la sentencia `library(Rcmdr)`.

#### ***d. Funciones básicas de R-commander.***



## Opción **Fichero**



Una de las primeras opciones que debemos conocer es el Directorio de Trabajo, que no es más que la carpeta del ordenador en la que el programa trabajará por defecto (bien para leer datos, bien para guardar algún resultado, gráfica,...). El directorio de trabajo se puede establecer en la opción de menú **Fichero > Cambiar directorio de trabajo**. En la ventana que aparece indicaremos en qué carpeta del ordenador vamos a trabajar.

Todo el trabajo que realicemos en R commander a través de las diferentes opciones de menú, el programa lo irá transformando en instrucciones que aparecerán en la **Ventana de instrucciones**. Es posible guardar todas las instrucciones generadas en un fichero para poder repetir en otro momento exactamente el mismo trabajo a través de la opción de menú: **Fichero > Guardar las instrucciones como**. O bien cargar el fichero de instrucciones existente (de un trabajo anterior, por ejemplo), mediante la opción de menú: **Fichero > Abrir fichero de instrucciones**.

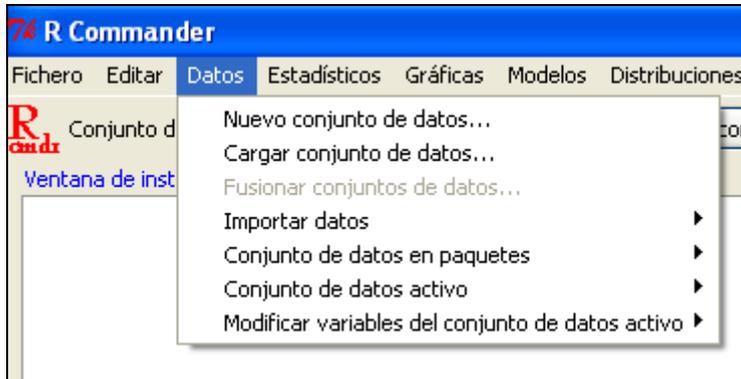
En general funcionaremos con las opciones de menú de R commander, pero si con el tiempo, aprendemos que determinadas acciones se corresponden con una determinada instrucción podremos escribir esta instrucción directamente en la **Ventana de instrucciones** sin necesidad de acceder a las opciones de menú en cuestión.



CEU

Universidad  
Cardenal Herrera

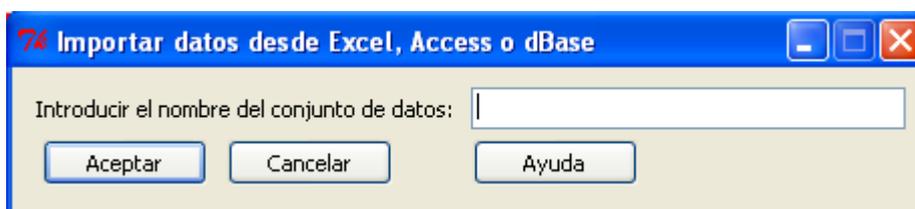
## Opción **Datos**



En la opción de menú **Datos**, se nos da acceso a la creación de un fichero de datos o bien a la importación de ficheros existentes, bien de ficheros creados en R con anterioridad, o bien de ficheros de datos de otras plataformas (Excel, Access,...). Es recomendable introducir los datos en un programa externo como Excel, Access,... cuyas funcionalidades permiten un manejo sencillo de los datos, y proceder a su importación al programa R-Commander únicamente para realizar el análisis.

Para ilustrar el funcionamiento del programa en la carga de datos y modificación de variables supongamos que estamos interesados en importar el fichero de datos [Alumnos.xls](#), y sobre este fichero realizaremos algunas acciones que pasamos a detallar.

Para importar un fichero de datos "Access" o "Excel", por ejemplo, accedemos al menú **Datos** y elegimos la opción **Importar datos** y a continuación **desde conjunto de datos Access, Excel o dBase...**, y el programa nos pedirá qué nombre queremos darle al fichero de datos que vamos a cargar:



Introducimos el nombre que queramos y pulsamos **Aceptar**. A continuación se nos abrirá una ventana de exploración en la que tendremos que localizar el fichero deseado, seleccionarlo y aceptar (En el caso de ficheros de Excel se nos pedirá qué hoja del fichero queremos cargar, y en el caso de ficheros de Access se nos pedirá que tabla de ese fichero queremos cargar).

Notar que los alumnos que dispongan de un ordenador MAC no tendrán la opción de importar datos Excel/Access, por lo que ponemos a disposición de estos alumnos un fichero llamado [Alumnos.rda](#) que pueden cargar

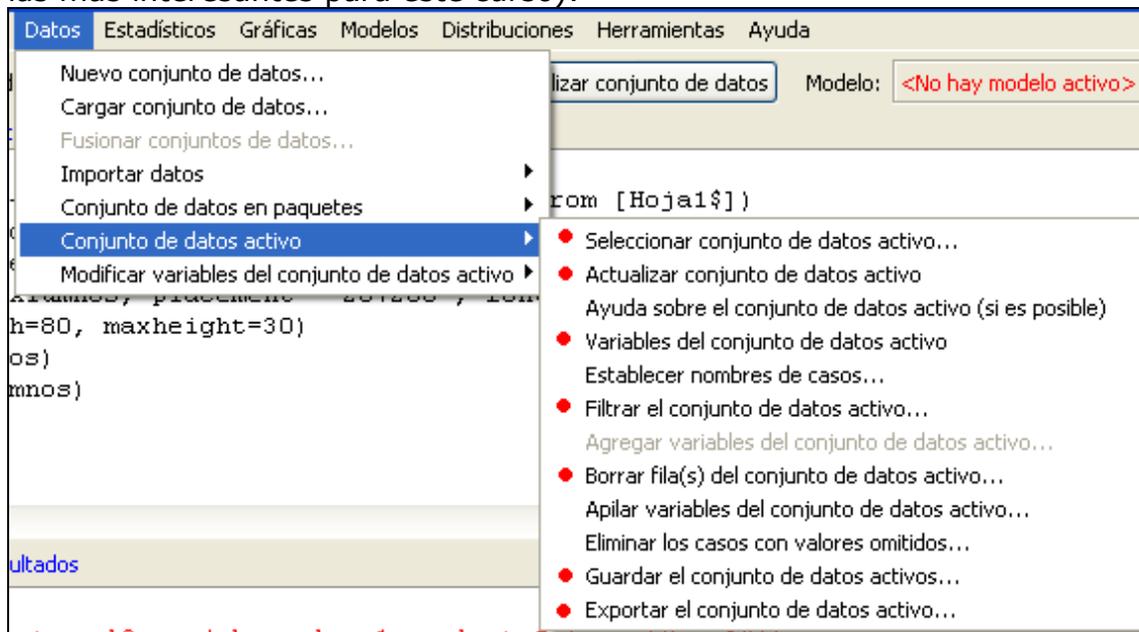
directamente desde la opción **Datos>Cargar datos** para que puedan reproducir toda la funcionalidad que se detalla a continuación.

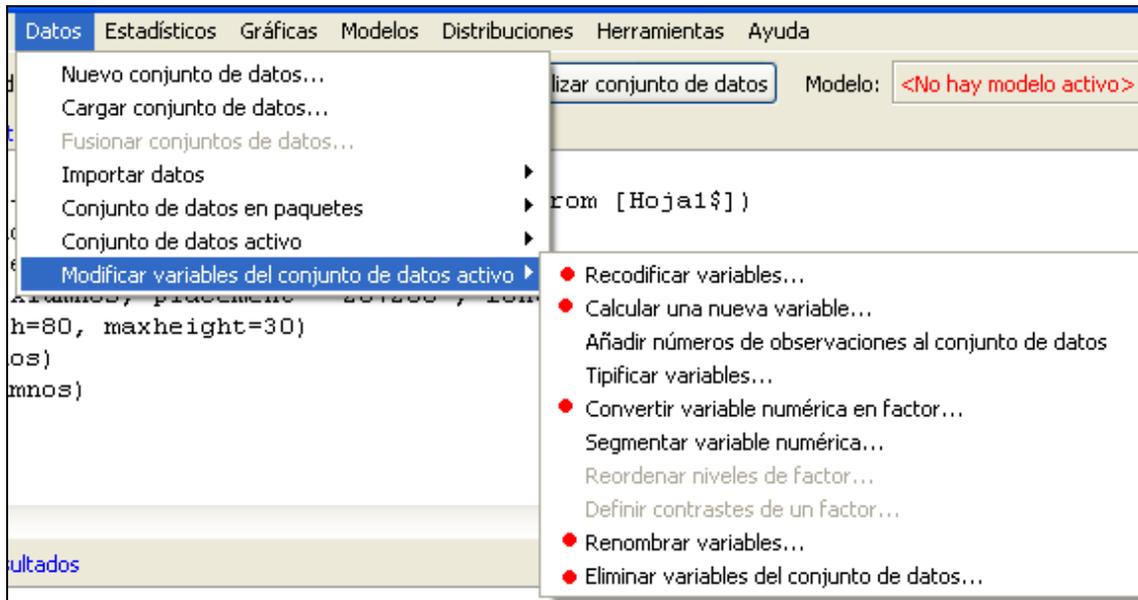
Observaremos en la **Ventana de instrucciones** el comando que el programa ha ejecutado para cargar los datos (y también en la **Ventana de resultados**).

Si queremos visualizar o modificar el fichero de datos cargado podemos hacerlo seleccionando en la ventana principal de **R commander** las opciones **Editar conjunto de datos** o bien **Visualizar conjunto de datos**. Podemos observar en la parte superior izquierda del programa, qué fichero de datos tenemos cargado en ese momento:



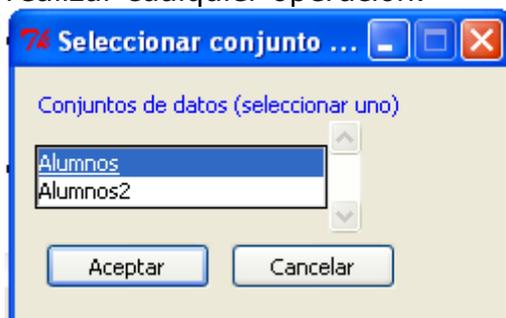
Para realizar operaciones con los registros del fichero de datos cargados, puedes acceder al menú de R commander Datos, a continuación o bien **Conjunto de datos activo** o bien **Modificar variables del conjunto de datos activo**. A continuación se muestran gráficamente las sub-opciones de cada una de estas dos opciones (Hemos marcado con un punto rojo • las más interesantes para este curso):





Breve descripción de las funciones resaltadas:  
(Opción **Conjunto de datos activo**)

- **Seleccionar conjunto de datos activo...:** Esta opción permite seleccionar, de entre todos los conjuntos de datos cargados, cuál es el "activo" y por tanto cuál es al banco de datos sobre el que vamos a realizar cualquier operación.

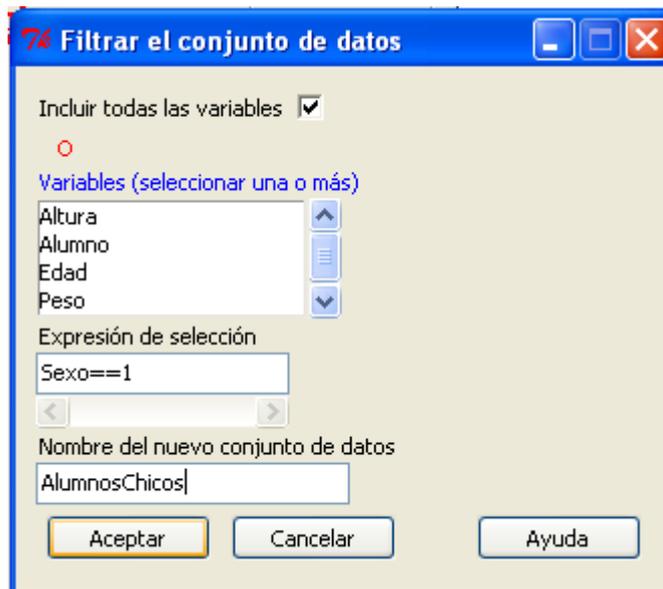


- **Actualizar conjunto de datos activo:** Si se realiza cualquier operación de modificación del conjunto de datos en origen, con esta opción se actualizan las modificaciones en el fichero cargado en R.
- **Variables del conjunto de datos activo:** Esta opción devuelve el nombre de las variables recogidas en el conjunto de datos (nombre de cada una de las columnas).
- **Filtrar el conjunto de datos activo:** Permite filtrar el conjunto de datos activo seleccionando únicamente las filas (registros) que cumplan una condición. Por ejemplo, si nos interesara seleccionar únicamente los alumnos "chicos" para hacer un subanálisis con ellos y en la variable Sexo se indica 1 si el alumno es "chico" y 2 si el alumno es "chica", podríamos obtener un nuevo conjunto de datos que contuviera únicamente a los "chicos" (al nuevo fichero le llamamos "AlumnosChicos" indicando en la ventana de filtro:



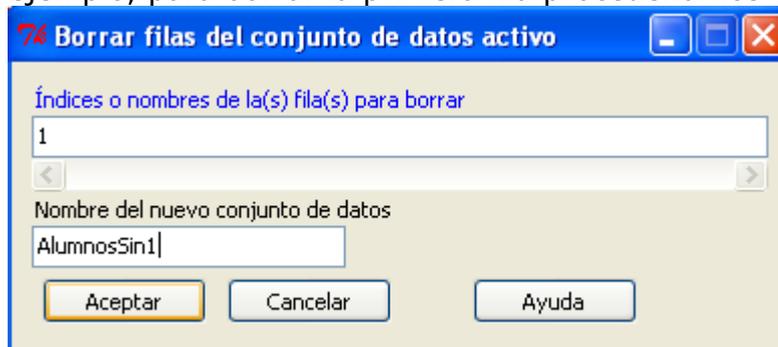
CEU

Universidad  
Cardenal Herrera



Para indicar que una variable sea igual a un valor indicamos "`==valor`", para menor, menor o igual, mayor o mayor o igual indicamos respectivamente "`<valor`", "`<=valor`", "`>=valor`" y "`>valor`".

- **Borrar fila(s) del conjunto de datos activo:** Mediante esta opción podemos borrar alguna de las filas del conjunto de datos activo. Por ejemplo, para borrar la primera fila procederíamos indicando:



- **Guardar el conjunto de datos activo:** Mediante esta opción, R Commander nos permite guardar el fichero de datos activo en un fichero de R, con extensión ".rda", en cualquier directorio del ordenador.
- **Exportar el conjunto de datos activo:** Mediante esta opción, el programa permite exportar el conjunto de datos activo a un fichero ".txt" y guardarlo en cualquier directorio del ordenador.

(Opción **Modificar variables del conjunto de datos activo**)

- **Recodificar variables:** Esta opción es útil, cuando por ejemplo queremos crear una nueva variable en función de los valores de otra.



CEU

Universidad  
Cardenal Herrera

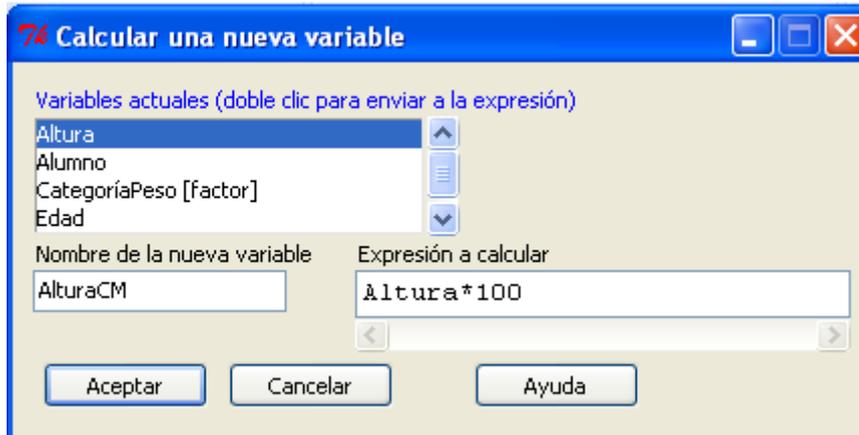
Supongamos que queremos recodificar la variable Peso, y transformarla en otra de tres categorías "Bajo", "Normal" o "Alto". En este caso podríamos proceder de la siguiente forma:



Esta opción crearía una nueva variable llamada "CategoríaPeso" que estaría codificada con el criterio establecido:

Peso	CategoríaPeso
65	Normal
55	Normal
59	Normal
67	Alto
62	Normal
70	Alto

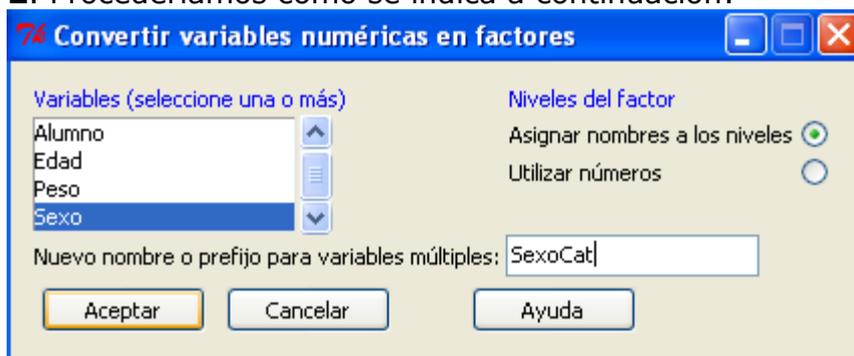
- **Calcular una nueva variable:** Esta opción es útil para las situaciones en las que queremos calcular una variable a partir de otra. Por ejemplo, podemos estar interesados en calcular la variable "Altura" en "cm", en lugar de en "m" (como está en el ejemplo). Así, procederíamos de la siguiente forma:



El programa nos crearía una nueva columna llamada "AlturaCM" en la que aparecerían los valores de la columna "Altura" multiplicados por 100:

Altura	AlturaCM
1.66	166
1.56	156
1.57	157
1.74	174
1.65	165
1.71	171

- **Convertir variable numérica en factor:** Esta opción permite convertir una variable codificada numéricamente a una variable categórica, e incluso asignar nombres a las categorías. Por ejemplo, para que el programa entienda que **Sexo** es una variable categórica podemos crear una nueva en la que aparezca "Hombre" cuando en Sexo aparezca el valor **1** y "Mujer" cuando en Sexo aparezca el valor **2**. Procederíamos como se indica a continuación:



Se crearía una nueva variable llamada **SexoCat**. Además, tendríamos que indicar los nombres que queremos para las dos categorías:



CEU

Universidad  
Cardenal Herrera



- **Renombrar variables:** Esta opción permite cambiar el nombre a cualquiera de las variables del conjunto de datos activo.
- **Eliminar variables del conjunto de datos...:** Esta opción permite eliminar variables (columnas) del conjunto de datos activo.

- **IMPORTANTE:** Por tanto, cuando quieras trabajar con R-Commander debes seguir los siguientes pasos:
- Abrir el programa **R** (bien desde el menú de Windows o bien desde el icono del escritorio).
  - Debes teclear en la ventana que aparece: **library(Rcmdr)**. A continuación se abrirá la ventana del programa.
  - Para cargar los datos debes considerar lo siguiente:
    - Si los datos que quieres cargar son externos al programa R, por ejemplo de Excel, Access, etc., debes acudir a la opción **Datos>Importar Datos**.
    - Si los datos que quieres cargar son del programa R, porque has trabajado en él anteriormente y los has guardado o porque te lo ha proporcionado el profesor, el fichero tendrá en su nombre las extensiones "XXX.rda" o "XXX.rdata". En ese caso debes acudir a la opción **Datos>Cargar Datos**.



CEU

Universidad  
Cardenal Herrera

## 2. Estadística descriptiva con R-Commander

### a. *Introducción.*

Para ilustrar el funcionamiento de R-Commander en estas técnicas, trabajaremos con un banco de datos sobre pacientes con migraña. Está compuesto por 90 individuos que asisten a un especialista en neurología para tratar sus migrañas. Se dispone de tres tratamientos ("A", "B" y "C"), y cada paciente ha sido asignado al azar a uno de los tres. Las variables recogidas para cada paciente son:

- **Sexo:** Codificado como 1 (Hombre) o 2 (Mujer).
- **Altura:** Medida en cm
- **Peso:** Medido en kg
- **HbAntes:** Nivel de hemoglobina en la última analítica, antes de tratar la migraña farmacológicamente.
- **FM:** Frecuencia de las migrañas, codificada como "Menos de 1 vez por semana" o "1 vez o más por semana".
- **HbDespues:** Nivel de hemoglobina después de tratar la migraña farmacológicamente.
- **AnteFami:** Codificado como 1 (tiene antecedentes familiares con el mismo problema de migrañas) y 0 (no tiene antecedentes familiares).
- **trata:** Tratamiento que ha recibido en sesiones presenciales ("A", "B" o "C")
- **tiempo** : Tiempo (en días) que ha pasado sin migrañas tras recibir el tratamiento

**gradodolor:** Grado de dolor en las sesiones del tratamiento medido como "0=dolor muy intenso", "1=dolor intenso", "2=dolor suave", "3=dolor muy suave".



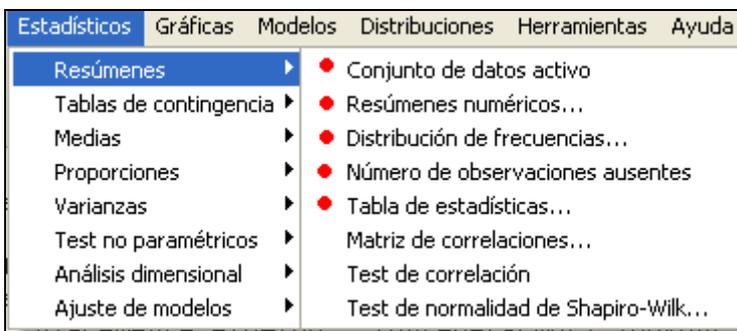
	Sexo	Altura	Peso	Hb	FM
1	1	177.04	77.34	12.61	Menos 1 vez por semana
2	2	176.47	67.16	12.35	Menos 1 vez por semana
3	2	152.71	59.52	12.46	1 vez o más por semana
4	2	155.58	52.6	11.49	Menos 1 vez por semana
5	1	163.09	77.19	12.27	1 vez o más por semana
6	2	159.73	53.95	12.73	1 vez o más por semana
7	1	162.65	68.93	12.63	Menos 1 vez por semana
8	2	147.45	52.83	12.52	Menos 1 vez por semana
9	1	179.91	81.6	13.57	Menos 1 vez por semana
10	1	161.54	71.81	12.99	Menos 1 vez por semana
11	2	173.8	58.65	12.9	Menos 1 vez por semana

Supongamos que este fichero se ha cargado en R Commander con el nombre **Datos.rda**. Para cargarlo y para poder trabajar con él y reproducir el análisis descriptivo que se va a explicar a continuación, carga el fichero "Datos.Rda" con la opción de menú **Datos>Cargar conjunto de datos...**

A continuación comentaremos algunas opciones de R Commander para realizar estadística descriptiva para una variable (cuantitativa o cualitativa) y estadística descriptiva bivariable, tanto numérica como gráfica.

## ***b. Resumen numérico de variables cuantitativas y cualitativas***

Exploraremos las opciones de menú que cuelgan de la opción principal **Estadísticos**, a continuación seleccionamos **Resúmenes**, y del desplegable estudiaremos las que aparecen en la siguiente figura marcadas con un punto rojo ● en la siguiente figura.



### **Opción Conjunto de datos activo**

Esta opción realiza un resumen general de cada una de las variables del banco de datos. Calcula los siguientes estadísticos:

<b>Variables cuantitativas</b>	<b>Variables cualitativas</b>	<b>categóricas</b>	<b>o</b>
<b>Frecuencias absolutas</b>	Mínimo ( <b>Min.</b> )		
	Primer cuartil ( <b>1st Qu.</b> )		
	Mediana ( <b>Median</b> )		
	Media ( <b>Mean</b> )		
	Tercer cuartil ( <b>3rd Qu.</b> )		
	Máximo ( <b>Max.</b> )		

Mostraría en la ventana de resultados:

```
> summary(Datos)
Sexo      Altura      Peso      Hb      FM
1:30  Min.   :118.8  Min.   :49.34  Min.   :11.48  1 vez o más por semana:13
2:20  1st Qu.:160.3  1st Qu.:58.23  1st Qu.:12.39  Menos 1 vez por semana:37
      Median :165.1  Median :62.91  Median :12.67
      Mean   :164.9  Mean   :63.55  Mean   :12.66
      3rd Qu.:176.6  3rd Qu.:68.34  3rd Qu.:12.94
      Max.   :200.7  Max.   :81.60  Max.   :13.89
```



CEU

Universidad  
Cardenal Herrera

### Opción **Resúmenes numéricos...**

En esta opción, se resumen una o varias variables cuantitativas simultáneamente:

Seleccionamos las variables que queremos resumir. Pulsando CTRL podemos seleccionar más de una.

Por defecto calculará la MEDIA, DESVIACIÓN TÍPICA y los CUANTILES que indica (MÍNIMO, PRIMER CUARTIL, MEDIANA, TERCER CUARTIL y MÁXIMO). Si queremos omitir alguno de los estadísticos o cambiar los cuantiles podemos cambiar la selección por defecto.

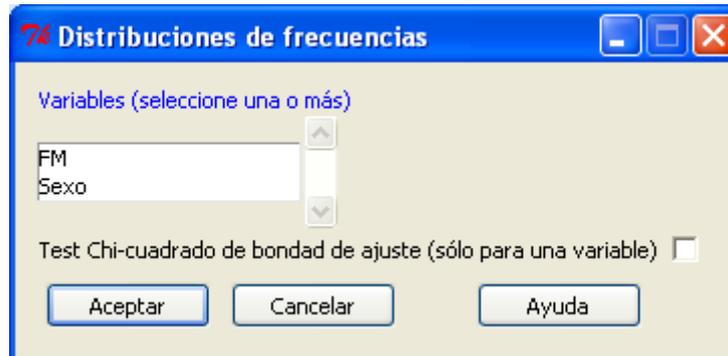
Si con la selección por defecto, seleccionamos todas las variables (únicamente ofrece las cuantitativas) y pulsamos aceptar el programa mostrará de la siguiente forma los cálculos realizados:

	mean	sd	0%	25%	50%	75%	100%	n
Altura	164.9144	14.8862949	118.78	160.270	165.090	176.5975	200.69	50
Hb	12.6568	0.4815131	11.48	12.395	12.670	12.9375	13.89	50
Peso	63.5474	7.4989626	49.34	58.230	62.915	68.3450	81.60	50

(n indica el número de observaciones para cada variable)

### Opción **Distribución de frecuencias...**

Mediante esta opción, se ofrece la posibilidad de resumir, mediante las frecuencias relativas, las variables categóricas. El programa nos mostrará una ventana en la que podremos elegir la/s variable/s (sólo mostrará las categóricas). Podremos seleccionar una de ellas, o varias pulsando CTRL después de la primera:



Tras seleccionar la/s variable/s deseada el programa mostrará los resultados en la Ventana de Resultados:

```
> .Table # counts for FM
1 vez o más por semana Menos 1 vez por semana
          13                      37
> 100*.Table/sum(.Table) # percentages for FM
1 vez o más por semana Menos 1 vez por semana
          26                      74
```

Frecuencias absolutas

Frecuencias relativas

### Opción **Número de observaciones ausentes**

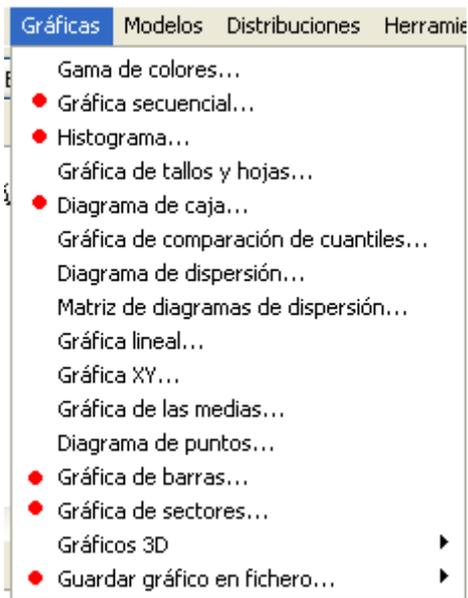
En esta opción, el programa muestra el número de valores ausentes en cada una de las variables del banco de datos activo.

```
Sexo  Altura  Peso  Hb  FM
    0     0     0     0     0
```

Indica que ninguna de las variables en el ejemplo tiene ningún valor ausente

## c. Resumen gráfico de variables cuantitativas o categóricas

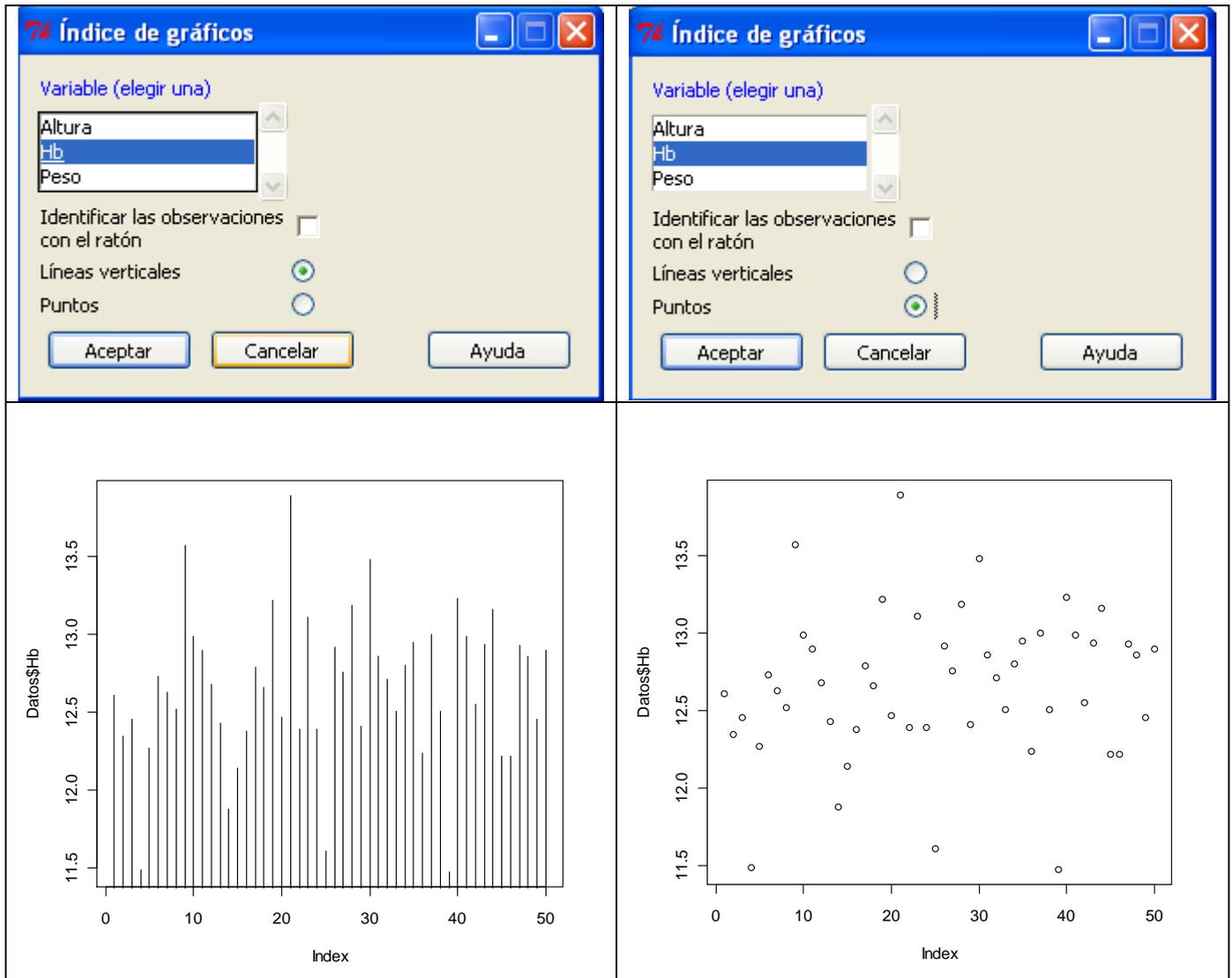
Exploraremos las opciones de menú que cuelgan de la opción principal **Gráficas**, concretamente las que aparecen en la siguiente figura marcadas con un punto rojo ●.



### Opción **Gráfica secuencial...**

Esta opción de menú, resulta especialmente útil, cuando tratamos datos secuenciales, por ejemplo en el tiempo. Supongamos que hemos recogido, el número de clientes por día atendidos en una farmacia, o el nivel de Hb de una embarazada para cada semana del embarazo, mediante esta opción visualizaríamos gráficamente el comportamiento de esta variable a través del tiempo.

En el ejemplo que estamos trabajando, no hay ninguna variable de este tipo, pero podemos observar el tipo de gráfico si, por ejemplo, representamos el nivel de Hb de los pacientes analizados:



(Existe la opción de representar los valores mediante **Puntos** o mediante **Líneas verticales**)

En cualquiera de las opciones gráficas, podemos observar, tanto en la Ventana de Instrucciones, como en la Ventana de resultados las instrucción correspondiente. Por ejemplo, en el caso del gráfico secuencial mediante puntos la instrucción ha sido:

```
plot(Datos$Hb, type="p")
```

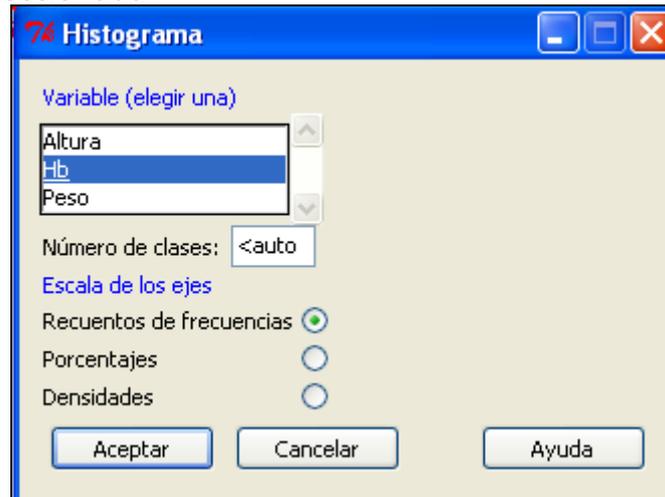
Si escribimos en la ventana de instrucciones la misma instrucción, pero añadimos **col="red"**, repetirá el mismo gráfico representando en esta ocasión los puntos en color "rojo". Así, si escribimos en la Ventana de instrucciones la siguiente expresión `plot(Datos$Hb, type="p", col="red")` y pulsamos **Ejecutar**, obtendremos el mismo gráfico pero en distinto color.



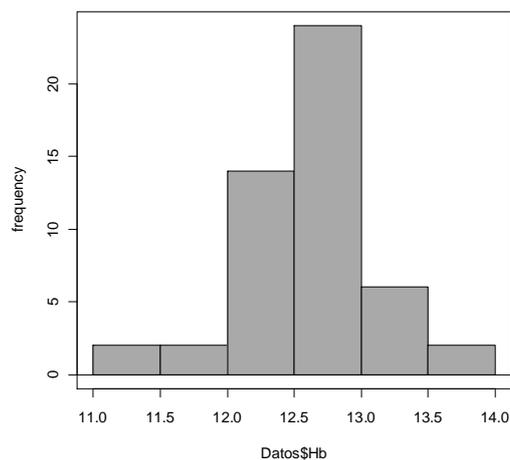
La opción "col" es una opción disponible en la mayoría de funciones gráficas, que permite cambiar el color de los gráficos que se van a realizar.

### Opción **Histograma...**

Mediante esta opción, el programa realiza un histograma de la variable cuantitativa seleccionada:



Nos permite elegir la cantidad a representar, o bien recuentos de frecuencias, o bien porcentajes. La forma del histograma, que es en principio lo que nos interesa, no varía.



De nuevo en esta opción, podríamos cambiar la apariencia del histograma cambiando el color. Para ello escribiríamos en la ventana de instrucciones la misma instrucción que se muestra cuando pedimos un histograma, pero añadiríamos la opción "col", con el color deseado.

Existen otras opciones interesantes que pueden variar la apariencia de los gráficos que se muestran. De entre todas estas opciones, merece la pena



CEU

Universidad  
Cardenal Herrera

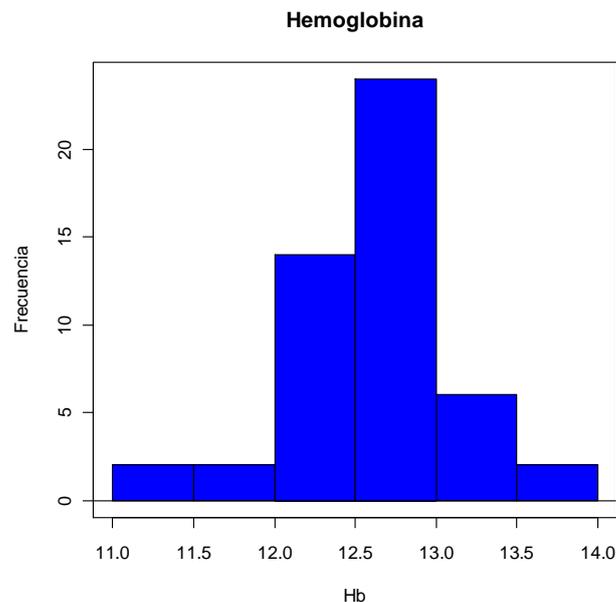
resaltar las "etiquetas" que el gráfico muestra en los ejes y el título del mismo.

Si por ejemplo, en el caso del histograma que acabamos de mostrar, la instrucción que da el programa es:

```
Hist(Datos$Hb, scale="frequency", breaks="Sturges", col="darkgray")
```

podríamos cambiar las etiquetas de los ejes y añadir un título modificando esta instrucción y añadiendo "xlab" (para la etiqueta del eje x), "ylab" (para la etiqueta del eje y) y "main" para el título:

```
Hist(Datos$Hb, scale="frequency", breaks="Sturges", col="blue",xlab="Hb",ylab="Frecuencia",main="Hemoglobina")
```



**En las últimas versiones de R-Commander** existe la posibilidad de incluir las etiquetas de los ejes y el título en una pestaña de "Opciones" del gráfico, en la que se nos muestran una sección similar a la siguiente en la que podemos añadir lo que deseemos para estas etiquetas:

Dibujar etiquetas

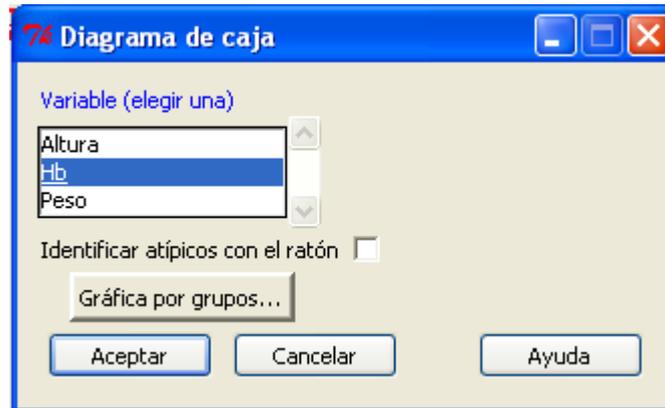
Etiqueta del eje x <auto>

Etiqueta del eje y <auto>

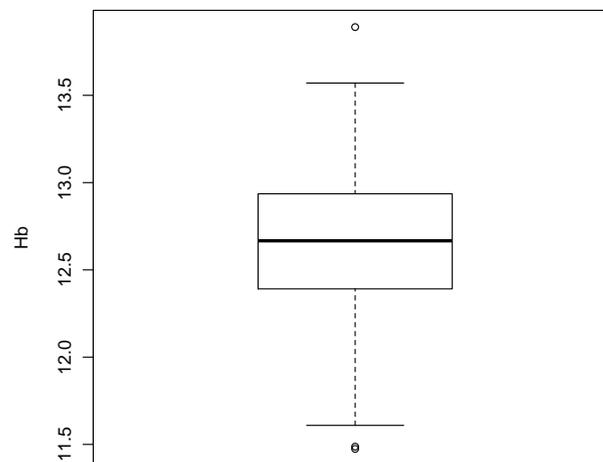
Título del gráfico <auto>

### Opción **Diagrama de caja...**

Mediante esta opción, el programa presenta un diagrama de caja de la variable cuantitativa seleccionada:

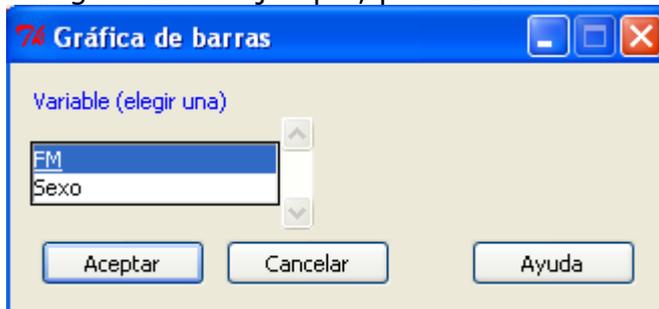


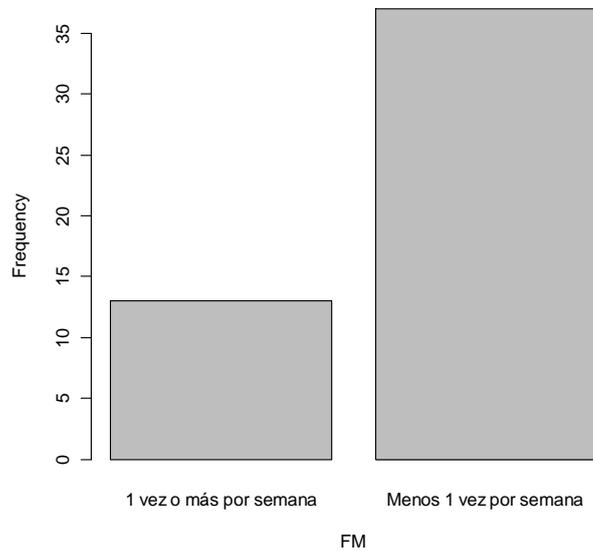
Por ejemplo, para la variable "Hb" mostraría:



### Opción **Gráfica de barras...**

Esta opción muestra una gráfica de barras para representar una variable categórica. Por ejemplo, para la variable "FM" se obtendría:





### Opción **Gráfica de sectores...**

Esta opción, al igual que la opción anterior, se utiliza para representar variables categóricas. En el caso de nuevo de la variable "FM" obtendríamos:

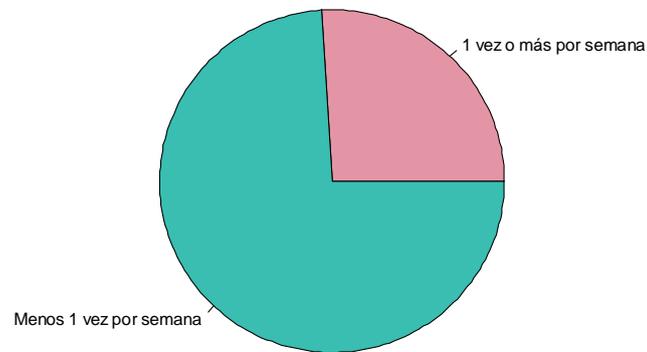




CEU

Universidad  
Cardenal Herrera

FM



### Opción **Guardar gráfico en fichero...**

Mediante esta opción el programa nos permite guardar cualquiera de los gráficos realizados en cualquier formato.

Además, pinchando con el ratón sobre cualquier gráfico y pulsando el botón derecho también podemos guardar el gráfico o copiarlo para pegarlo posteriormente sobre un documento, por ejemplo.

## d. Resumen numérico/gráfico bivariable

### Variables: Una cuantitativa y otra cualitativa:

#### Opción **Estadísticos>Resúmenes > Resúmenes numéricos...**

En esta opción, se resumen una o varias variables cuantitativas simultáneamente, como se ha visto anteriormente. Mediante la opción **Resumir por grupos...**, podemos realizar los cálculos comentados para las diferentes categorías de una variable categórica. Si, por ejemplo, queremos el resumen de las variables cuantitativas Altura, Hb y Peso para "Hombre" por un lado y para "Mujeres" por otro, seleccionaremos las tres variables cuantitativas como antes y en la opción **Resumir por grupos... Sexo:**



El programa mostrará, entonces, los siguientes resultados:

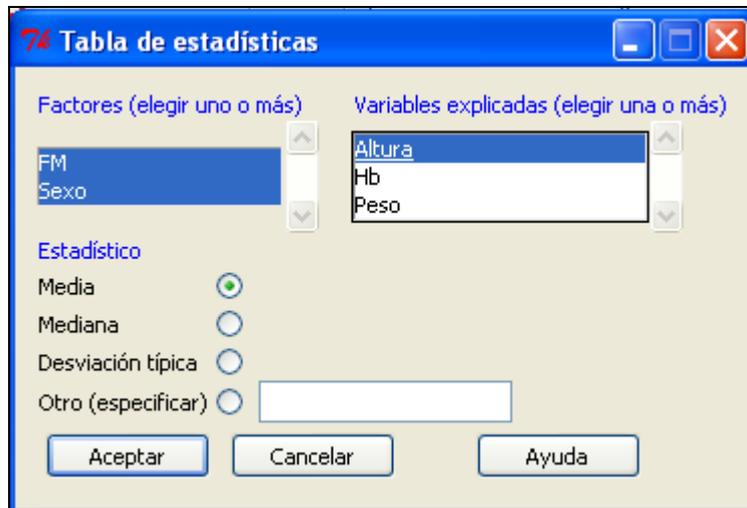
Variable: Altura								
	mean	sd	0%	25%	50%	75%	100%	n
1	167.8650	14.62194	118.78	161.675	166.100	177.64	200.69	30
2	160.4885	14.51914	130.08	151.395	163.225	171.01	182.83	20
Variable: Hb								
	mean	sd	0%	25%	50%	75%	100%	n
1	12.64933	0.4225288	11.48	12.395	12.67	12.915	13.57	30
2	12.66800	0.5702871	11.49	12.410	12.64	13.040	13.89	20
Variable: Peso								
	mean	sd	0%	25%	50%	75%	100%	n
1	66.91867	7.314866	50.28	61.4100	67.455	72.2750	81.60	30
2	58.49050	4.300362	49.34	55.8875	58.915	60.6725	67.16	20

(Recordemos que **1** indica **Hombre** y **2** indica **Mujer**)

#### Opción **Estadísticos>Resúmenes>Tabla de estadísticas**

La opción **Tabla de estadísticas**, permite calcular estadísticos de variables cuantitativas (media, mediana, desviación típica) para las categorías de una variable categórica o el cruce de varias variables categóricas. Si en el ejemplo queremos calcular la altura media de los hombres con más o

menos frecuencia de migrañas, seleccionaremos por un lado la variable altura, y por otra las dos categóricas (Sexo y FM):



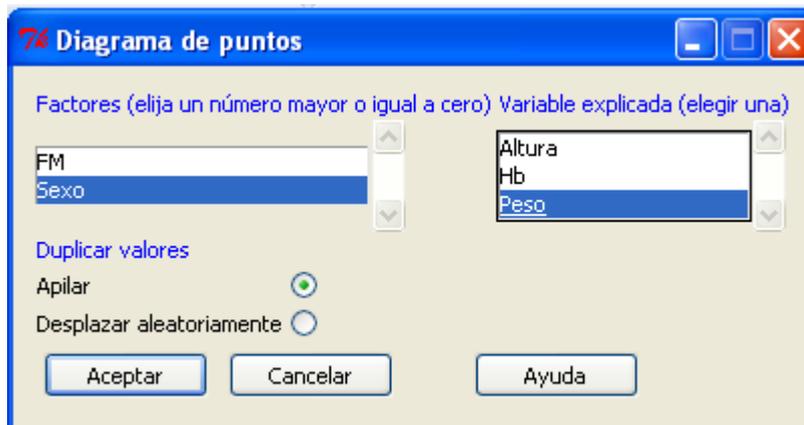
El programa mostrará el resultado en una tabla (altura media para el cruce de las dos variables categóricas):

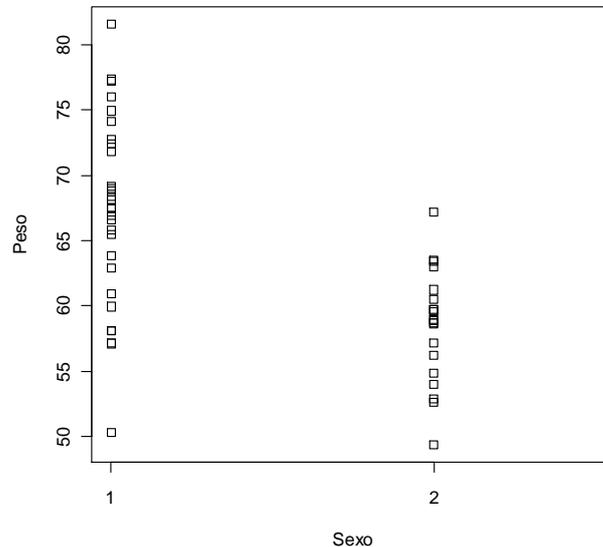
		Sexo	
FM		1	2
1 vez o más por semana		168.8844	157.9825
Menos 1 vez por semana		167.4281	161.1150

### Opción **Gráficas>Diagrama de puntos...**

En esta opción se muestra la dispersión (mediante un punto para cada dato) de los valores de una variable cuantitativa para cada una de las categorías de una variable cualitativa.

Por ejemplo, si mostramos la variable "Peso", para los grupos que define la variable "Sexo":

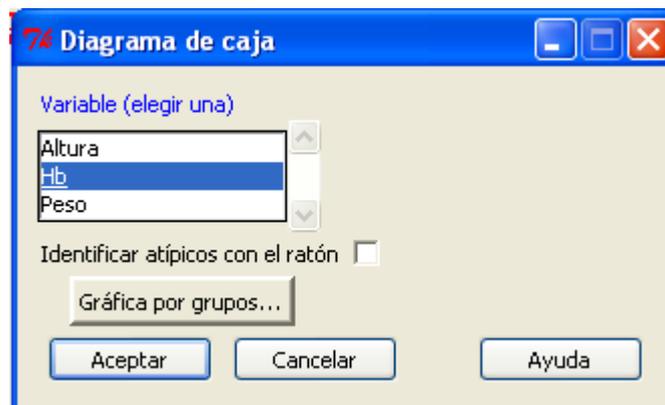




Podemos observar que los valores de 1 (Hombres) son mayores que los de 2 (Mujeres) y que además muestran más dispersión.

### Opción **Gráficas>Diagrama de caja...**

Mediante esta opción, el programa presenta un diagrama de caja de la variable cuantitativa seleccionada y en la opción **Gráfica por grupos...**, nos permite seleccionar una variable cualitativa y mostrará un diagrama de caja para cada grupo definido por esa variable cualitativa. Por ejemplo, para la variable cuantitativa Hb y seleccionando en Gráfica por grupos "Sexo" obtendríamos una diagrama de caja para "Hombres" y otro para "Mujeres":

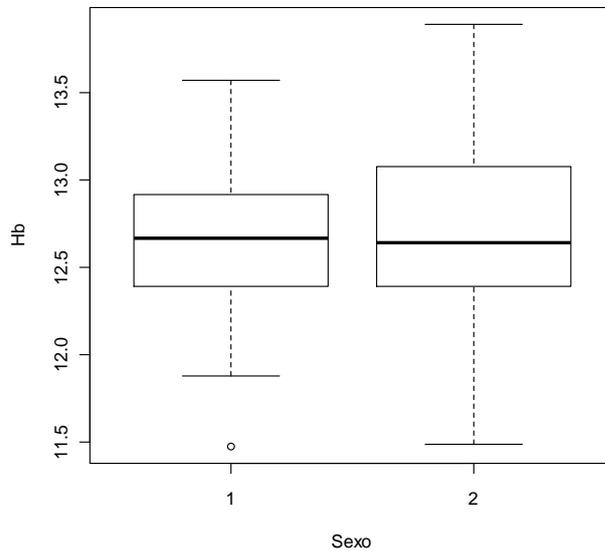


Por ejemplo, para la variable "Hb" mostraría:



CEU

Universidad  
Cardenal Herrera



## **Variables: Dos variables cuantitativas**

### **Opción Gráficas>Diagrama de dispersión...**

Esta opción contiene muchísimas posibilidades, pero a esta altura del temario, la utilizaremos para poder representar conjuntamente, mediante una nube de puntos, dos variables categóricas. Así, si por ejemplo queremos comprobar la relación existente entre las variables "Peso" y "Altura", seleccionaríamos ambas variables, y **sin seleccionar ninguna otra opción** el programa mostraría el comportamiento conjunto de ambas variables:



**74 Diagrama de dispersión**

variable x (elegir una) variable y (elegir una)

Altura Hb Peso Altura Hb Peso

**Options**

Identificar Observaciones

Desplazar aleatoriamente la variable x

Desplazar aleatoriamente la variable y

Log eje-x

Log eje-y

Cajas de dispersión marginales

Línea de mínimos cuadrados

Línea suavizada

Amplitud para suavizado 50

**Dibujar parámetros**

Dibujar los caracteres <auto>

Tamaño del punto 1.0

Tamaño del texto (ejes) 1.0

Tamaño del texto (etiquetas de los ejes) 1.0

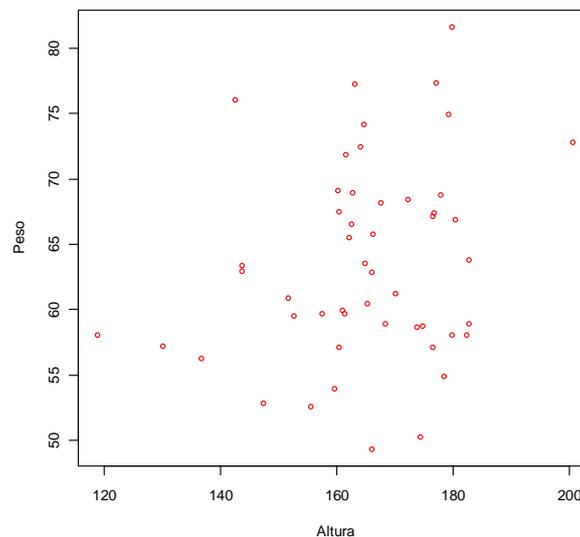
Etiqueta del eje x <auto>

Etiqueta del eje y <auto>

Expresión de selección <todos los casos válidos>

Gráfica por grupos...

Aceptar Cancelar Ayuda

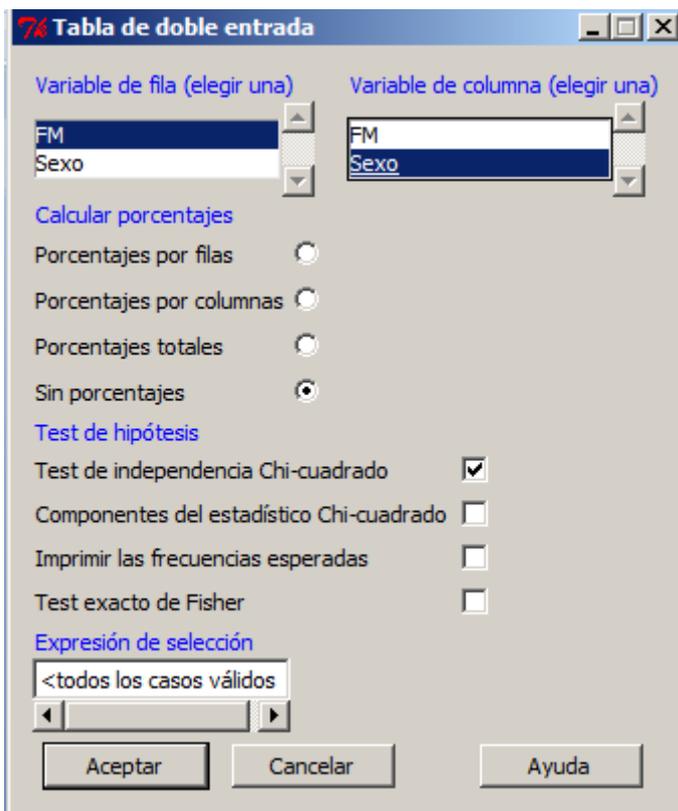


Si quisiéramos estimar la posible relación lineal entre dos variables, con la opción **Línea de mínimos cuadrados**, se representaría gráficamente esta relación lineal. En la sesión de *Regresión lineal* se profundizará en este análisis.

## **Variables: Dos variables cualitativas**

**Opción Estadísticos>Tablas de contingencia>Tablas de doble entrada...**

Esta opción nos permite visualizar (además de otras cuestiones que se abordarán en la sesión de Inferencia) en una tabla de contingencia la relación de dos variables cualitativas. Accediendo a esta opción y seleccionando las dos variables que deseamos resumir:



El programa nos muestra la siguiente salida:

```
Ventana de resultados
> .Table <- xtabs(~FM+Sexo, data=Datos)
> .Table
              Sexo
FM
 1 vez o más por semana  9  4
Menos 1 vez por semana 21 16
```

En esta ventana podemos ver cómo se reparten los individuos en las diferentes categorías que generan el cruce de las dos variables y analizar si nos parece que hay o no relación entre ambas.