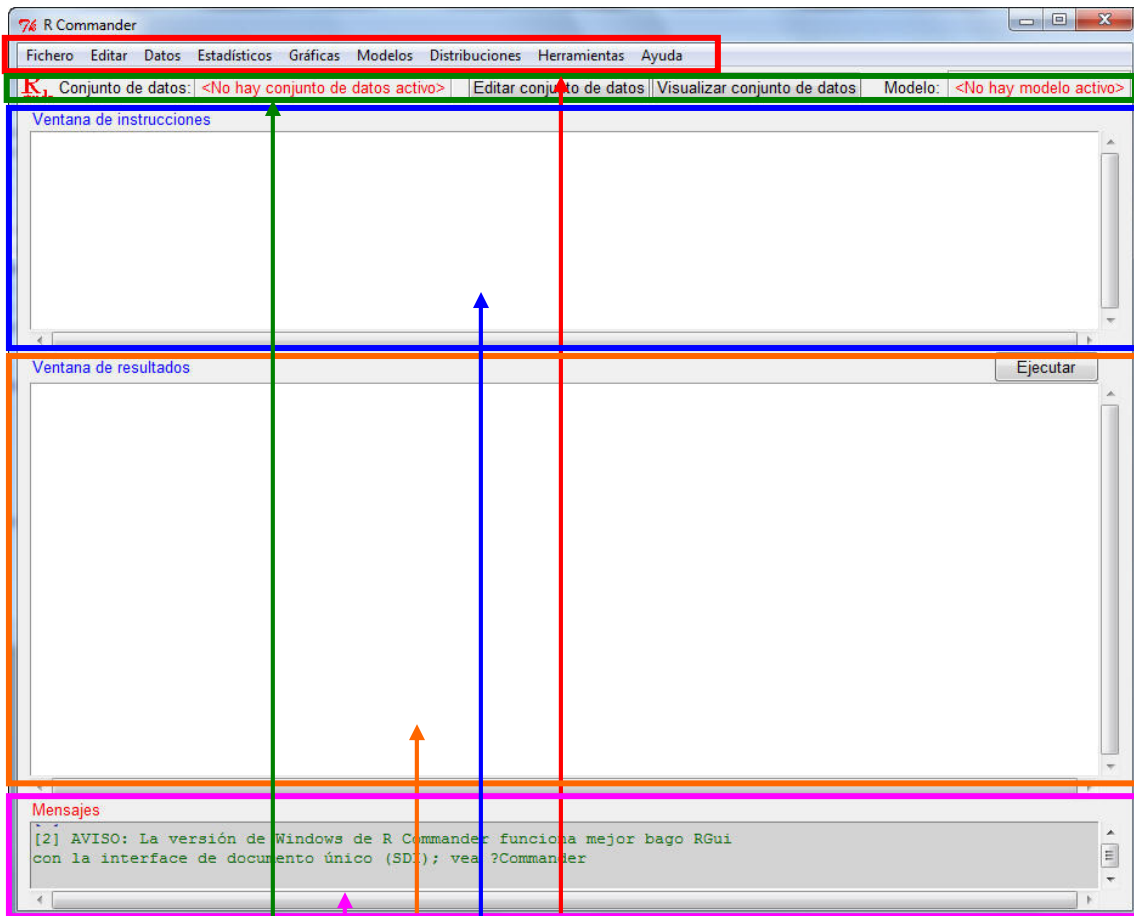


Comenzando con R Commander

Una vez cargado R Commander, aparecerá la siguiente ventana



en la que podemos distinguir varias partes:

- Un menú de ventanas desplegable: Fichero, Editar, Datos, Estadísticos, ... Algunas de ellas permiten gestionar ficheros y otras contienen menús con los análisis estadísticos disponibles.
- Una cabecera con ventanas asociadas al conjunto de datos y modelo en activo.
- Una ventana de instrucciones en la que aparecerá el código de R asociado con cada acción que ejecutemos desde un menú. Como ya comentamos, esto permite familiarizarnos con dicho código y poder repetir alguna acción (o variante de ella, retocando el código) sin tener que volver a utilizar el menú. Además, esta ventana equivale a la consola de R y, por tanto, podremos ejecutar una instrucción directamente sin más que escribirla, marcarla y clicar en el botón *Ejecutar*.
- Una ventana de resultados, donde aparecen los resultados asociados a cualquier operación o análisis realizado.
- Una ventana de mensajes, en la que aparece cualquier información asociada con la ejecución de alguna orden, en particular, mensajes de error.

Nota importante: Si se cierra R Commander sin cerrar R, para volver a cargarlo se debe ejecutar la instrucción

> Commander()

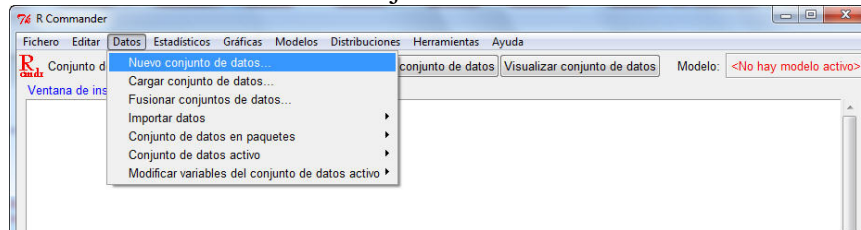
Manejo de datos con R Commander

Introducción de un fichero de datos

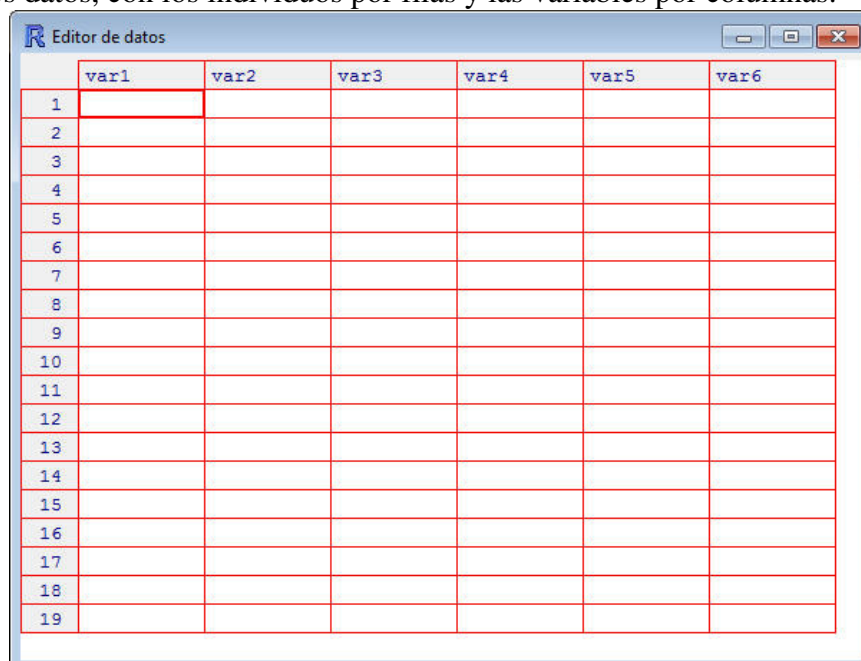
Seleccionamos

Datos

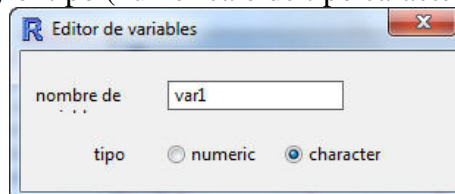
Nuevo conjunto de datos ...



y nos solicitará el nombre de dicho conjunto de datos. Una vez introducido un nombre, pulsamos *Aceptar*. A continuación se abre una ventana con casillas en la que debemos escribir los datos, con los individuos por filas y las variables por columnas.



Clicando con el ratón en la parte superior de cada columna, podemos introducir el nombre de cada variable y el tipo (numérica o de tipo carácter).



Para terminar, cerramos la ventana del editor de datos y R Commander almacena dicho datos y los considera como *Conjunto de datos activo* mostrando el nombre del fichero en la cabecera asociada al conjunto de datos y modelo en activo.

Una vez introducido o cargado un *conjunto de datos* se puede editar o visualizar clicando sobre el botón correspondiente.

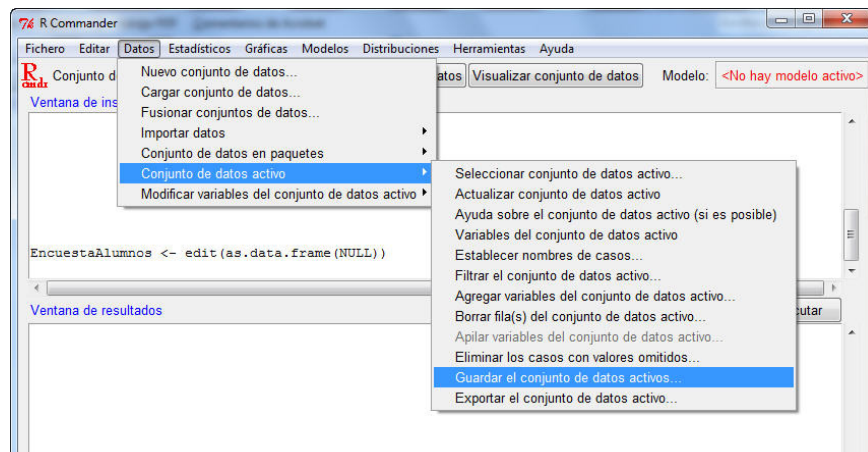
Almacenamiento de un fichero de datos

El fichero de datos que se introduce (ver sección anterior) se almacena de forma temporal para trabajar con él pero si cerramos R se eliminará. Para poder usarlo en una sesión de trabajo posterior debemos guardarlo. Para ello, seleccionamos

Datos

Conjunto de datos activo

Guardar el conjunto de datos activos



Nos pedirá un nombre (con extensión .RData) y un directorio donde lo almacenará.

Si posteriormente, queremos cargar ese fichero de datos, se seleccionará

Datos

Cargar conjunto de datos ...

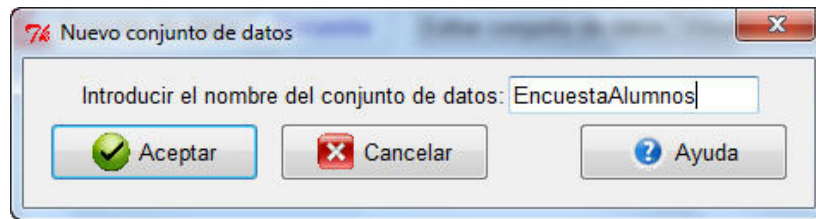
y se buscará el archivo en la ventana del explorador que se muestra.

Ejemplo

Vamos a introducir los datos resultantes de preguntar a cinco estudiantes del Grado en Matemáticas el curso en que están matriculados y las horas diarias que dedican al estudio:

Curso	Horas
Primero	3
Segundo	2
Primero	1
Tercero	4
Segundo	2

Llamamos al fichero *EncuestaAlumnos*

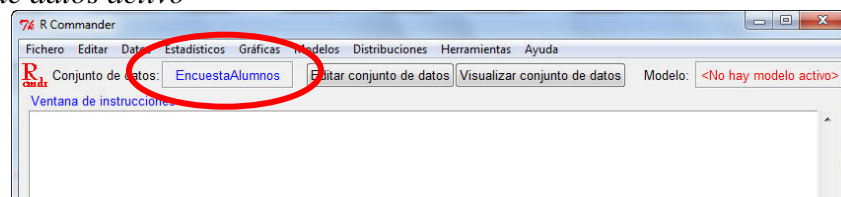


Llamamos a la primera variable *Curso*, que será de tipo “cualitativo” y a la segunda *Horas*, que será de tipo “numérica”.

Una vez introducido los datos, se debe mostrar la siguiente pantalla

	Curso	Horas	var3	var4	var5	var6
1	Primero	3				
2	Segundo	2				
3	Primero	1				
4	Tercero	4				
5	Segundo	2				
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						

Si cerramos el editor de datos, R Commander los almacena y los muestra como *Conjunto de datos activo*

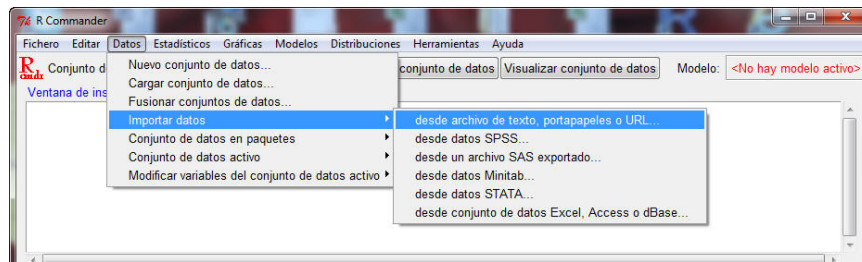


Recordemos que si queremos usarlo en una sesión de trabajo posterior, se debe guardar con extensión *.RData*.

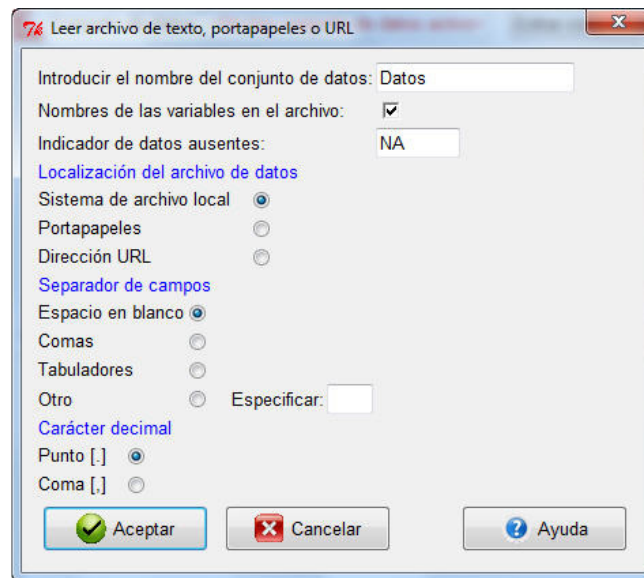
Cómo importar un fichero de datos

Seleccionamos

Datos
Importar datos



Si consideramos la opción “desde archivo de texto, portapapeles o URL...” se abre la siguiente ventana

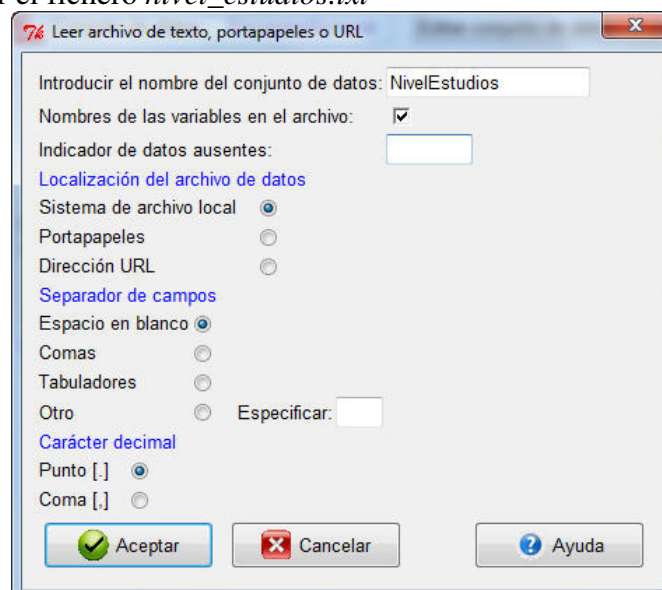


donde debemos especificar las opciones del fichero que queremos abrir. Si se selecciona en [Localización del archivo de datos](#) la opción “Sistema de archivo local”, debemos buscar el archivo en nuestro disco duro para la cual se abrirá una ventana del Explorador para que lo seleccionemos.

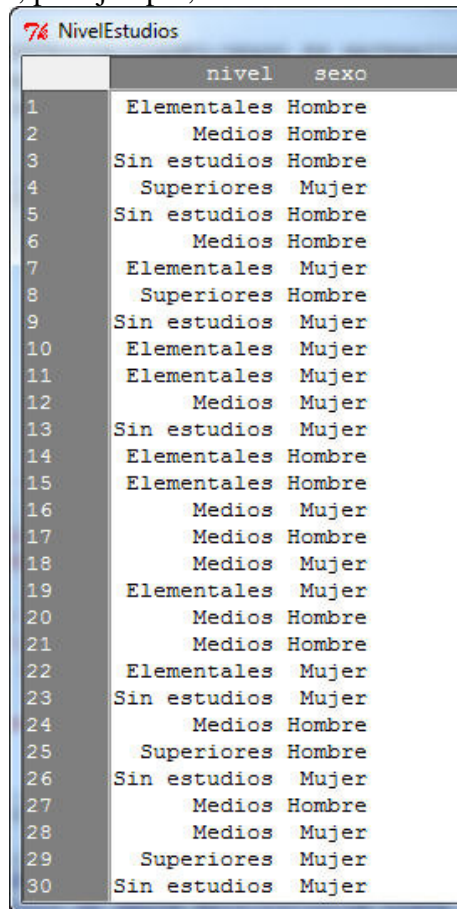
También se pueden importar datos de SPSS, SAS, Minitab, STATA, Excel, etc

Ejemplo

Vamos a importar el fichero *nivel_estudios.txt*



Una vez cargado podemos, por ejemplo, visualizarlo:



	nivel	sexo
1	Elementales	Hombre
2	Medios	Hombre
3	Sin estudios	Hombre
4	Superiores	Mujer
5	Sin estudios	Hombre
6	Medios	Hombre
7	Elementales	Mujer
8	Superiores	Hombre
9	Sin estudios	Mujer
10	Elementales	Mujer
11	Elementales	Mujer
12	Medios	Mujer
13	Sin estudios	Mujer
14	Elementales	Hombre
15	Elementales	Hombre
16	Medios	Mujer
17	Medios	Hombre
18	Medios	Mujer
19	Elementales	Mujer
20	Medios	Hombre
21	Medios	Hombre
22	Elementales	Mujer
23	Sin estudios	Mujer
24	Medios	Hombre
25	Superiores	Hombre
26	Sin estudios	Mujer
27	Medios	Hombre
28	Medios	Mujer
29	Superiores	Mujer
30	Sin estudios	Mujer

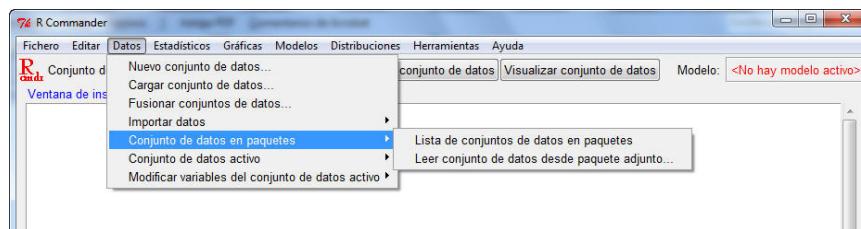
Ahora ya se puede trabajar con él y, si queremos, guardarlo con extensión .RData para poder usarlo en otra sesión de trabajo sin tener que volver a importarlo.

Cómo cargar conjuntos de datos desde paquetes

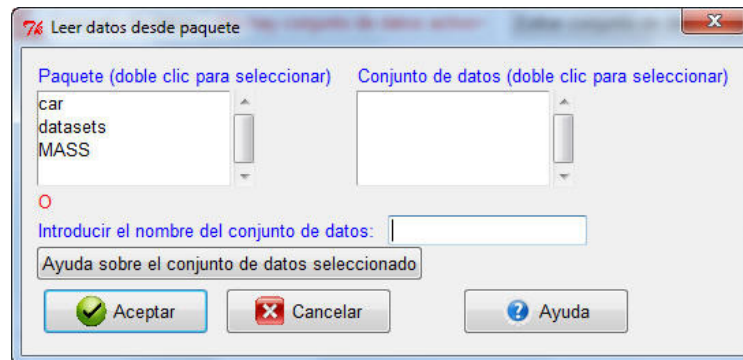
Existen paquetes de R con conjuntos de datos que pueden ser utilizados. Para cargar un conjunto de datos de un paquete concreto, seleccionamos

Datos

Conjunto de datos en paquetes



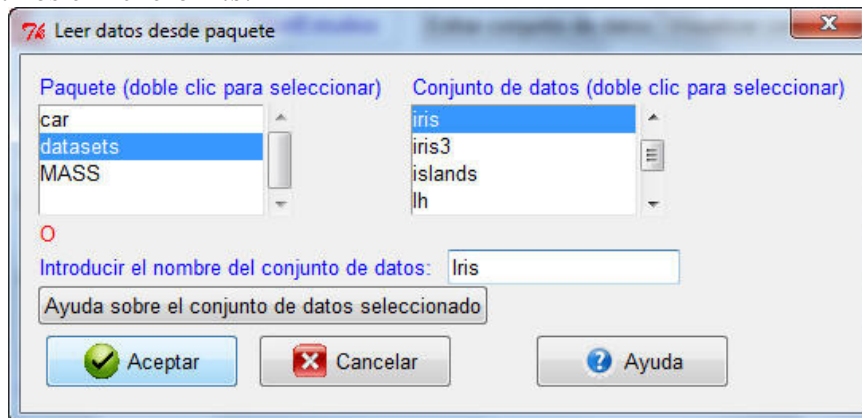
Si seleccionamos ahora “Lista de conjuntos de datos en paquetes” se muestra la lista de los paquetes disponibles. Si seleccionamos “Leer conjunto de datos desde paquete adjunto...” se abre la siguiente ventana



en la que debemos seleccionar, en primer lugar el paquete y, una vez seleccionado uno de ellos, un conjunto de datos, de la lista que aparecerá en la segunda ventana.

Ejemplo

Vamos a cargar el fichero *Iris* del paquete “datasets”. Hacemos doble click en el paquete “datasets” de la primera ventana de la figura anterior y se mostrarán en la segunda ventana todos los ficheros de datos contenidos en dicho paquete. De entre ellos, seleccionamos el fichero *Iris*:



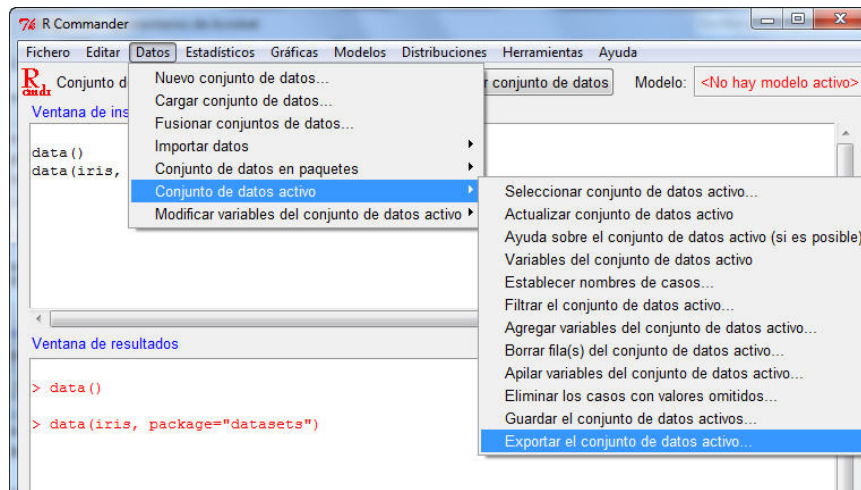
Cómo exportar un conjunto de datos

Ya hemos visto como guardar un conjunto de datos como fichero con formato de R (con extensión .RData). Si queremos exportar un conjunto de datos a formato .txt, .dat o .csv, seleccionamos

Datos

Conjunto de datos activo

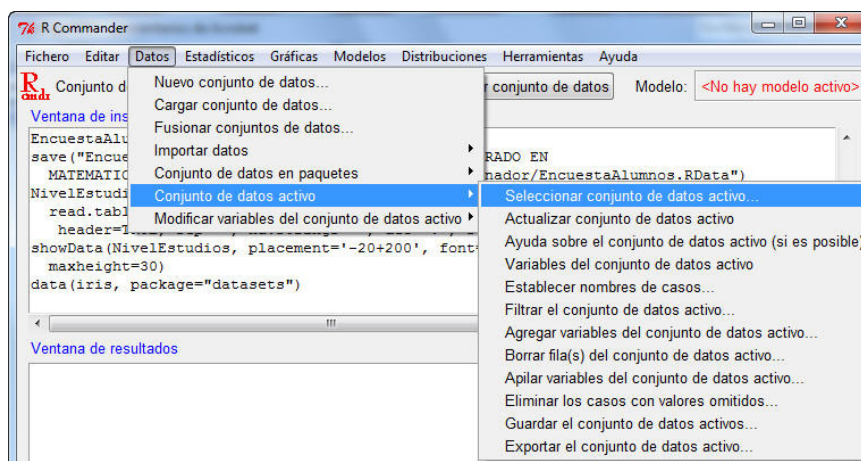
Exportar el conjunto de datos activo



Se abrirá una ventana donde debemos especificar las opciones del fichero que queremos exportar y, al darle a aceptar, una ventana del Explorador para que escribamos el nombre del archivo y seleccionemos la extensión y el directorio donde se guardará.

Ejemplo

Vamos a exportar el fichero de datos *EncuestaAlumnos* que habíamos creado antes. Al no haber cerrado R Commander está guardado aún y podemos recuperarlo como *Conjunto de datos activo* mediante



y seguimos los pasos indicados anteriormente para exportarlo.

Cómo manipular el conjunto de datos activo

Al seleccionar

Datos

Conjunto de datos activo

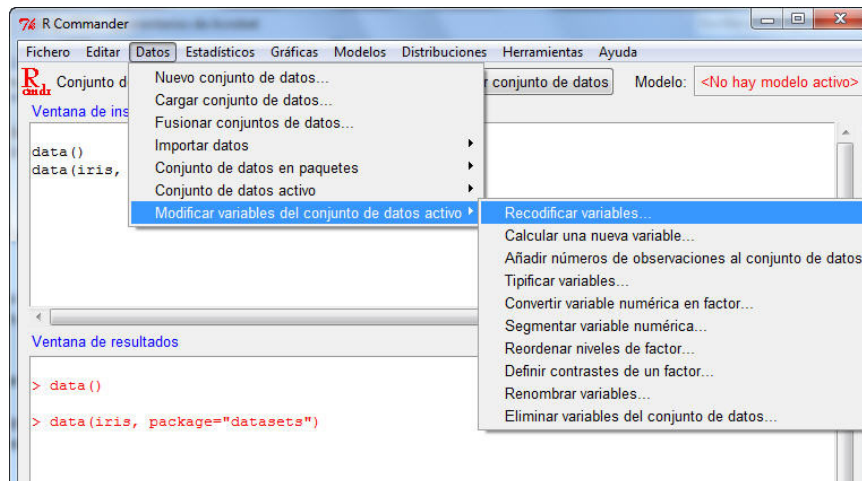
aparecen muchas opciones para permitir manipular el fichero de datos activo (filtrarlo, agregar variables, borrar filas, etc.).

Cómo modificar variables del conjunto de datos activo

Finalmente si queremos modificar variables, recodificar variables, calcular nuevas variables, segmentar variables numéricas, etc, del conjunto de datos activos, debemos seleccionar

Datos

Modificar variables del conjunto de datos activo



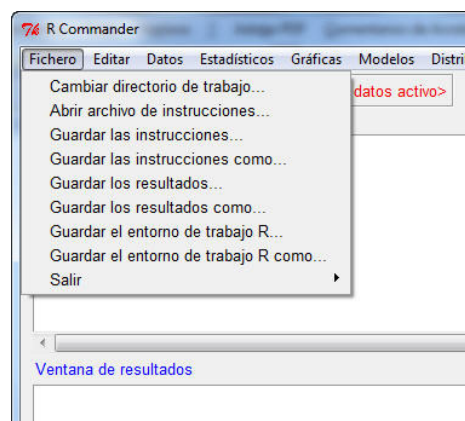
y luego una de las posibles opciones que se muestran. Analizaremos alguna de estas opciones más adelante en los ejemplos que se vayan realizando.

Almacenar instrucciones y resultados

Cuando se cierra una sesión de R Commander se nos pregunta por las opciones de

- Guardar el archivo de instrucciones (que se guardará con extensión .R). Luego se podrá recuperar y ejecutarlo de nuevo.
- Guardar el archivo de salida o de resultados (que se guardará con extensión .txt)

En cualquier momento también se pueden guardar dichos ficheros desde el menú "Fichero" que se muestra a continuación



Libros de R Commander

- **Estadística Básica con R y R-Commander**
Autores: A. J. Arriaza Gómez, F. Fernández Palacín, M. A. López Sánchez,
M. Muñoz Márquez, S. Pérez Plaza, A. Sánchez Navas.
Editorial: Servicio de publicaciones, Universidad de Cádiz

<http://knuth.uca.es/moodle/mod/resource/view.php?id=1126>

- **Métodos Estadísticos con R y RCommander**
Autor: Antonio Jesús Sáez Castillo.

<http://www4.ujaen.es/~ajsaez/recursos/RRCmdrv31.pdf>

Análisis estadístico unidimensional

Antes de comenzar con el análisis datos, vamos a ver cómo reconvertir a un fichero que R pueda tratar, datos que vienen dados en tablas de frecuencias.

Ejemplo 1

En una determinada empresa se realiza un estudio sobre la calidad de su producción. La distribución siguiente informa sobre el número de piezas defectuosas encontradas en 100 cajas examinadas con 50 unidades cada una de ellas:

Nº piezas defectuosas	0	1	2	3	4	5	6	7	8	9	10
Nº de cajas	6	9	10	11	14	16	16	9	4	3	2

Se deben incluir los datos en un fichero con extensión .dat (por ejemplo TablaPiezasDefectuosas.dat) de la siguiente forma

```
"PiezasDefect" "frec"
"1" 0 6
"2" 1 9
"3" 2 10
"4" 3 11
"5" 4 14
"6" 5 16
"7" 6 16
"8" 7 9
"9" 8 4
"10" 9 3
"11" 10 2
```

A continuación, cargamos ese fichero como

```
Datos
  Importar datos
    desde archivo de texto, portapapeles o URL...
```

Se le asocia como nombre al conjunto de datos “TablaPiezasDefectuosas”, por ejemplo, y para transformar la tabla de frecuencias en un conjunto de datos (data.frame) con el que R pueda trabajar hay que ejecutar las siguientes instrucciones

```
P<-rep(TablaPiezasDefectuosas$PiezasDefect,TablaPiezasDefectuosas$frec)
PiezasDefectuosas<-data.frame(P)
```

Por último, para trabajar con estos datos, debemos seleccionarlo como conjunto de datos activo (no olvidar exportarlo al final si queremos, en una sesión de trabajo posterior, volver a usarlo).

Si lo visualizamos queda

	PiezasDefect
1	0
2	0
3	0
4	0
5	0
6	0
7	1
8	1
9	1
10	1
11	1
12	1
13	1
14	1
15	1
16	2
17	2
18	2
19	2
20	2
21	2
22	2
23	2
24	2
25	2
26	3
27	3
28	3
29	3
30	3

Ejemplo 2

Nivel Estudios \ Sexo	Sin estudios	Elementales	Medios	Superiores
Hombre	432	670	1030	540
Mujer	450	600	910	450

Se deben incluir los datos en un fichero con extensión .dat (por ejemplo TablaFrecuencias.dat) de la siguiente forma

```
"nivel" "sexo" "frec"
"1" "Sin estudios" "Hombre" 432
"2" "Sin estudios" "Mujer" 450
"3" "Elementales" "Hombre" 670
"4" "Elementales" "Mujer" 600
"5" "Medios" "Hombre" 1030
"6" "Medios" "Mujer" 910
"7" "Superiores" "Hombre" 540
"8" "Superiores" "Mujer" 450
```

A continuación, cargamos ese fichero como

Datos

Importar datos

desde archivo de texto, portapapeles o URL...

Se le asocia como nombre al conjunto de datos `TablaFrec`, por ejemplo, y para transformar la tabla de frecuencias en un conjunto de datos (`data.frame`) con el que R pueda trabajar hay que ejecutar las siguientes instrucciones

```
nivel<-rep(TablaFrec$nivel, TablaFrec$frec)
sexo<-rep(TablaFrec$sexo, TablaFrec$frec)
NivelEstudios<-data.frame(nivel, sexo)
```

Por último, para trabajar con estos datos, debemos seleccionarlo como conjunto de datos activo.

Estadísticos

Resúmenes

Conjunto de datos activo

Mediante la orden anterior se obtiene un resumen de las variables del conjunto de datos activo. Para las variables cualitativas se muestra la distribución de frecuencias y para las variables numéricas los valores del mínimo, primer cuartil, mediana, media, tercer cuartil y máximo.

Ejemplo

Fichero *encues.txt* que contiene los resultados de una encuesta en la que se les ha preguntado a 50 personas su altura, nivel de estudios, número de personas que conviven en su domicilio habitual y su peso.

Después de haber importado el fichero y ejecutado la orden anterior, el resultado es el siguiente:

altura	estudios	habitantes	peso
Min. :1.340	bachiller:20	Min. :2	Min. : 33.50
1st Qu.:1.600	diplomado:10	1st Qu.:3	1st Qu.: 61.36
Median :1.730	fp : 5	Median :4	Median : 73.08
Mean :1.721	primario :13	Mean :4	Mean : 73.00
3rd Qu.:1.877	superior : 2	3rd Qu.:5	3rd Qu.: 88.35
Max. :2.020		Max. :8	Max. :105.01

Estadísticos

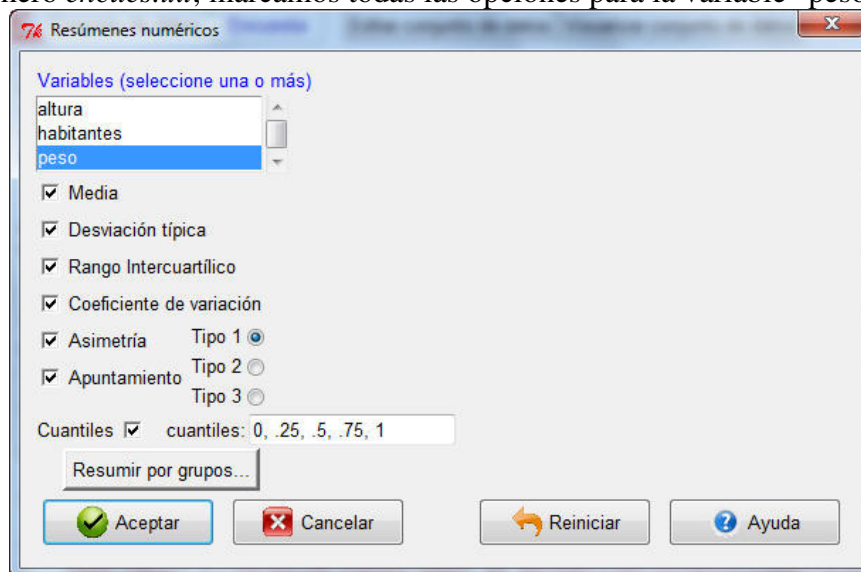
Resúmenes

Resúmenes numéricos

Al ejecutar esta secuencia de instrucciones se puede seleccionar, para cada una de las variables numéricas del conjunto de datos activo, el cálculo de la media, desviación típica, rango intercuartílico, coeficiente de variación, coeficiente de asimetría y de apuntamiento (en estos casos se debe seleccionar el tipo 1, para calcular los coeficientes de Fisher) y los cuantiles (pudiendo, en este caso, seleccionar los valores de órdenes deseados). Además, el estudio de cada variable numérica se puede realizar por grupos atendiendo a los valores de alguna variable cualitativa.

Ejemplo

Con el fichero *encues.txt*, marcamos todas las opciones para la variable “peso”



y el resultado es

mean	sd	IQR	cv	skewness	kurtosis
72.99704	18.98126	26.9925	0.2600278	-0.1911943	-0.7827098
0%	25%	50%	75%	100%	n
33.5	61.3575	73.075	88.35	105.01	50

Si seleccionamos la variable “altura” y en la ventana *Resumir por grupos...* seleccionamos la variable “estudios” el resultado es

	mean	sd	IQR	cv	skewness	kurtosis
bachiller	1.742000	0.18805374	0.3200	0.10795278	-2.031192e-01	-1.16408605
	0%	25%	50%	75%	100%	data:n
	1.38	1.5875	1.76	1.9075	2.02	20
	mean	sd	IQR	cv	skewness	kurtosis
diplomado	1.749000	0.13008117	0.1075	0.07437460	4.115194e-01	-0.03334798
	0%	25%	50%	75%	100%	data:n
	1.53	1.6800	1.74	1.7875	2.00	10
	mean	sd	IQR	cv	skewness	kurtosis
fp	1.634000	0.25095816	0.3400	0.15358517	8.928231e-02	-1.37413509

	0%	25%	50%	75%	100%	data:n
	1.34	1.4400	1.65	1.7800	1.96	5
primario	mean	sd	IQR	cv	skewness	kurtosis
	1.721538	0.20007691	0.1700	0.11621983	-3.208041e-01	-0.81268616
	0%	25%	50%	75%	100%	data:n
	1.36	1.6600	1.74	1.8300	1.99	13
superior	mean	sd	IQR	cv	skewness	kurtosis
	1.590000	0.07071068	0.0500	0.04447212	6.613633e-15	-2.00000000
	0%	25%	50%	75%	100%	data:n
	1.54	1.5650	1.59	1.6150	1.64	2

Como es lógico, al solicitar resúmenes numéricos sólo permite seleccionar variables cuantitativas.

Sin embargo a una variable cualitativa en escala ordinal se le puede someter a algún tipo de análisis numérico (por ejemplo, el cálculo de cuantiles) pero para ello debemos recodificarla numéricamente. Veamos cómo hacerlo con la variable “estudios” del fichero *nivel_estudios.txt*

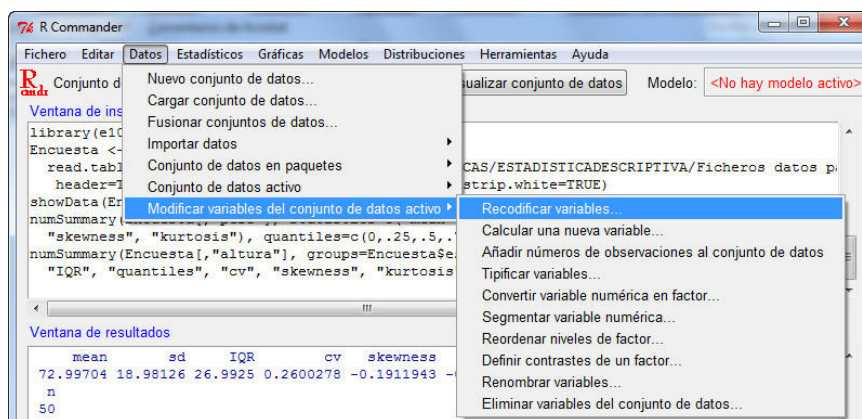
Tenemos dos opciones

1) Crear una nueva variable

Se selecciona

Datos

Modificar variables del conjunto de datos activo
Recodificar variables...



A continuación:

Se selecciona la variable

Se escribe un nuevo nombre

Se desmarca “Cada nueva variable en factor”

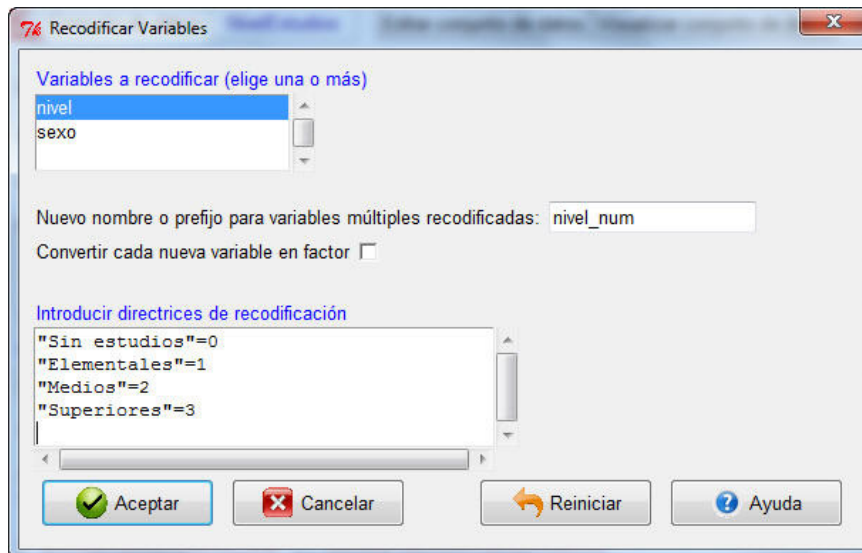
Se introducen los antiguos valores y los codificados de la siguiente manera

"Sin estudios"=0

"Elementales"=1

"Medios"=2

"Superiores"=3



Al solicitar

Estadísticos
Resúmenes
Resúmenes numéricos

ya se puede marcar la variable “nivel_num”

2) Reordenar los niveles de la variable

Antes de describir esta opción eliminamos la variable “nivel_num” mediante

Datos

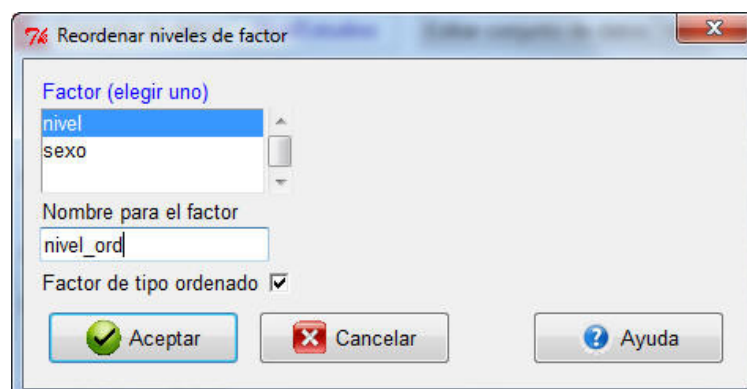
Modificar variables...
Eliminar variables ...

Ahora ejecutamos

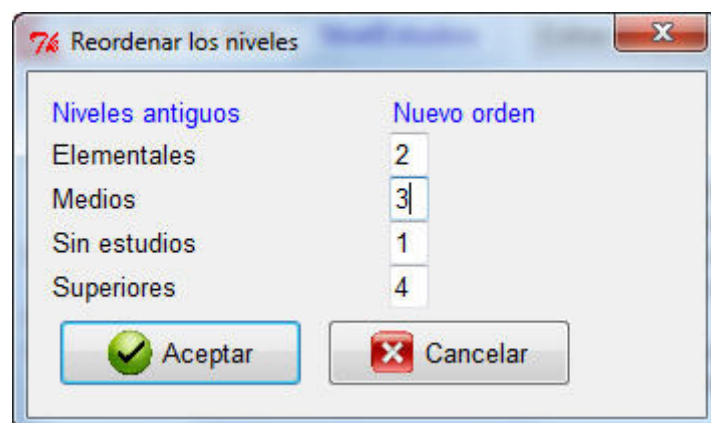
Datos

Modificar variables...
Reordenar niveles de factor ...

Seleccionamos la variable “nivel”, nombre para el factor “nivel_ord” y marcamos *Factor de tipo ordenado*.



Asignamos un valor numérico a cada modalidad, por ejemplo (notemos que, en esta opción, no admite el valor 0)



Finalmente, escribimos y ejecutamos

```
NivelEstudios$nivel_num<-as.numeric(NivelEstudios$nivel_ord)
```

Si visualizamos el nuevo conjunto de datos, se muestra

	nivel	sexo	nivel_ord	nivel_num
1	Elementales	Hombre	Elementales	2
2	Medios	Hombre	Medios	3
3	Sin estudios	Hombre	Sin estudios	1
4	Superiores	Mujer	Superiores	4
5	Sin estudios	Hombre	Sin estudios	1
6	Medios	Hombre	Medios	3
7	Elementales	Mujer	Elementales	2
8	Superiores	Hombre	Superiores	4
9	Sin estudios	Mujer	Sin estudios	1
10	Elementales	Mujer	Elementales	2
11	Elementales	Mujer	Elementales	2
12	Medios	Mujer	Medios	3
13	Sin estudios	Mujer	Sin estudios	1
14	Elementales	Hombre	Elementales	2
15	Elementales	Hombre	Elementales	2
16	Medios	Mujer	Medios	3
17	Medios	Hombre	Medios	3
18	Medios	Mujer	Medios	3
19	Elementales	Mujer	Elementales	2
20	Medios	Hombre	Medios	3
21	Medios	Hombre	Medios	3
22	Elementales	Mujer	Elementales	2
23	Sin estudios	Mujer	Sin estudios	1
24	Medios	Hombre	Medios	3
25	Superiores	Hombre	Superiores	4
26	Sin estudios	Mujer	Sin estudios	1
27	Medios	Hombre	Medios	3
28	Medios	Mujer	Medios	3
29	Superiores	Mujer	Superiores	4
30	Sin estudios	Mujer	Sin estudios	1

Estadísticos

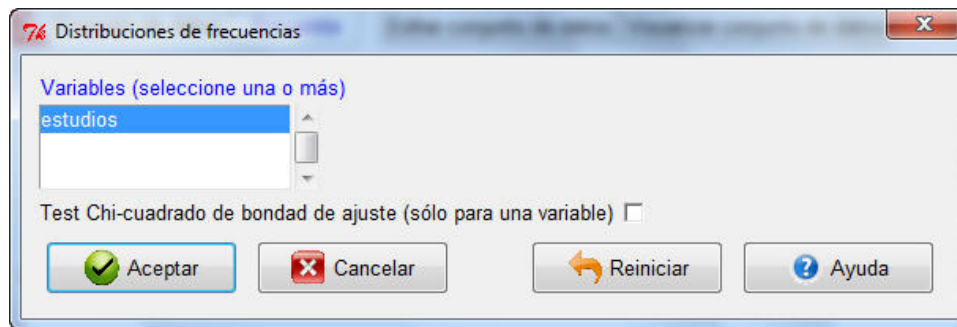
Resúmenes

Distribuciones de frecuencias

R Commander sólo permite, mediante menú, la realización de tablas de frecuencias para variables cualitativas.

Ejemplo

Si el fichero de datos activo es *Encuesta*, la única variable que permite seleccionar es “estudios”



El resultado es

```
> .Table <- table(Encuesta$estudios)
```

```
> .Table # counts for estudios
```

bachiller	diplomado	fp	primario	superior
20	10	5	13	2

```
> round(100* .Table/sum(.Table), 2) # percentages for estudios
```

bachiller	diplomado	fp	primario	superior
40	20	10	26	4

```
> remove(.Table)
```

* Si se quiere hacer una tabla de frecuencias para una variable numérica discreta lo tenemos que hacer mediante código de R. En general, los códigos son:

- Para distribución de frecuencias absolutas:
`table(Nombre fichero$NombreVariable)`

- Para número total de datos
`sum(table(Nombre fichero$NombreVariable))`

- Para frecuencias absolutas acumuladas
`cumsum(as.data.frame(table(Nombre fichero$NombreVariable))$Freq)`

- Para distribución de frecuencias relativas

```
table(Nombre fichero$NombreVariable)/sum(table(Nombre fichero$NombreVariable))
```

- Para distribución de porcentajes

```
100* table(Nombre fichero$NombreVariable)/sum(table(Nombre fichero$NombreVariable))
```

Ejemplo

Con el conjunto de datos *Encuesta*, la orden

```
table(Encuesta$habitantes)
```

proporciona

2	3	4	5	6	8
5	13	16	11	4	1

A continuación se muestran los resultados de la ejecución de las otras órdenes:

```
sum(table(Encuesta$habitantes))
```

50

```
cumsum(as.data.frame(table(Encuesta$habitantes))$Freq)
```

5 18 34 45 49 50

```
100* table(Encuesta$habitantes)/sum(table(Encuesta$habitantes))
```

2	3	4	5	6	8
10	26	32	22	8	2

* Si se quiere hacer una tabla de frecuencias para una variable numérica continua lo tenemos que hacer mediante código de R (lo veremos al hacer el histograma)

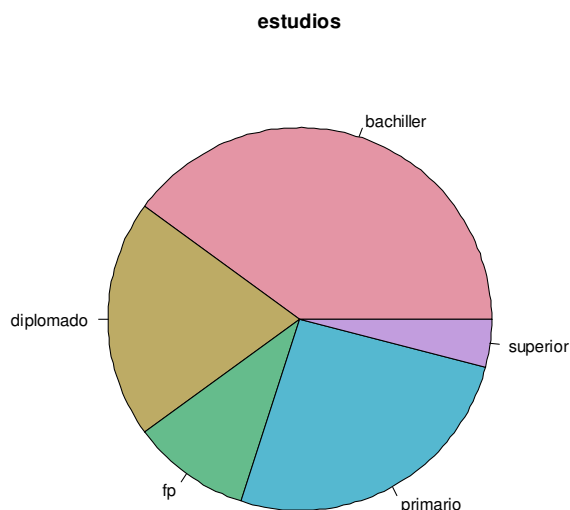
Gráficas

Gráfica de sectores

Sólo permite seleccionar variables de tipo cualitativo.

Ejemplo

Hacemos un gráfico de sectores con la variable “estudios” del conjunto de datos *Encuesta*. El resultado es

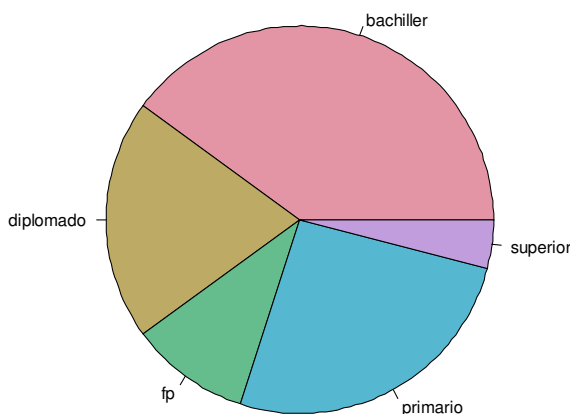


En la ventana de instrucciones, vemos que la instrucción que se ha ejecutado ha sido

```
pie(table(Encuesta$estudios), labels=levels(Encuesta$estudios), main="estudios",
  col=rainbow_hcl(length(levels(Encuesta$estudios))))
```

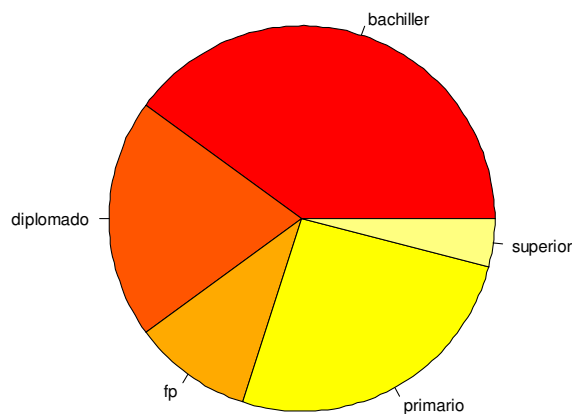
Podemos realizar cambiar en ese gráfico modificando el código. Por ejemplo, el título del gráfico se puede modificar en la opción main. Si cambiamos el texto “estudios” por “Diagrama de sectores para la variable Nivel de Estudios”, y volvemos a ejecutar la instrucción, el resultado es

Diagrama de sectores para la variable Nivel de Estudios



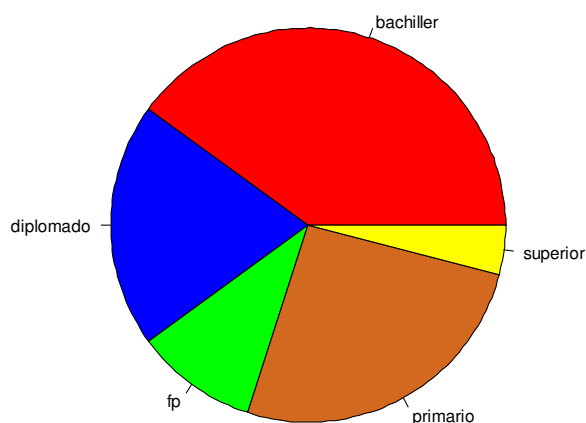
Otra posibilidad es cambiar la gama de colores. Por ejemplo, cambiar `col=rainbow_hcl(length(levels(Encuesta$estudios)))` por `col=heat.colors(5)`. (El valor 5 es debido a que hay 5 sectores). En este caso, el resultado es

Diagrama de sectores para la variable Nivel de Estudios



También podemos decir exactamente qué colores queremos usar (con nombres o en termino de sus components RGB con una cadena del tipo "#RRGGBB" donde cada par RR, GG, BB consiste en dos dígitos hexadecimales con valores en el rango de 00 a FF). Por ejemplo, con la opción `col=c("red","blue","green","chocolate","yellow")` se obtiene

Diagrama de sectores para la variable Nivel de Estudios



Una lista de nombres de colores se puede obtener con la orden `colors()` y en la web http://es.wikipedia.org/wiki/Colores_web aparece una lista de códigos RGB con los colores correspondientes.

Dos cuestiones relativas a gráficos en general son:

- Cómo guardar un gráfico: Una vez realizado un gráfico se puede guardar ejecutando

Gráficas

Guardar gráfico en archivo...

y seleccionado a continuación el formato, especificaciones, nombre y carpeta para guardarlo. Otra opción es, una vez posicionados en el gráfico, pulsar botón derecho del ratón y seleccionar la opción deseada.

- Cómo realizar un gráfico en una nueva ventana de forma que no se pierda otro anterior: Para ello antes de realizar un nuevo gráfico se debe ejecutar la orden `win.graph()`.

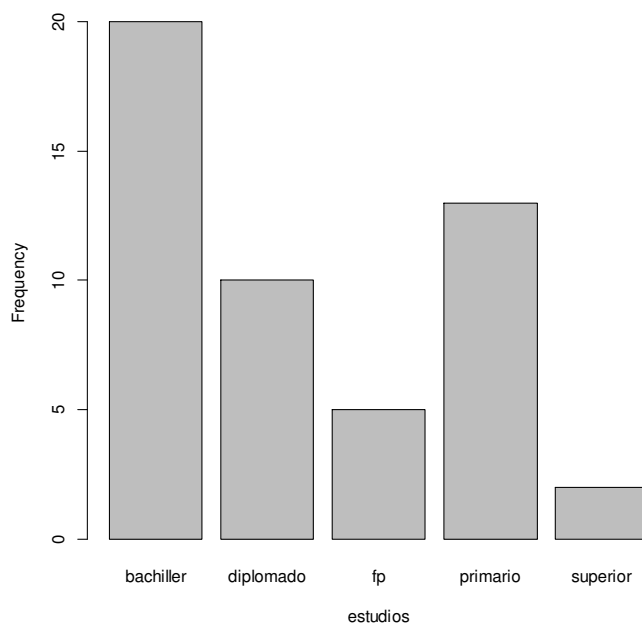
Gráficas

Gráfica de barras

Sólo permite seleccionar variables de tipo cualitativo.

Ejemplo

Hacemos un gráfico de sectores con la variable “estudios” del conjunto de datos *Encuesta*. El resultado es



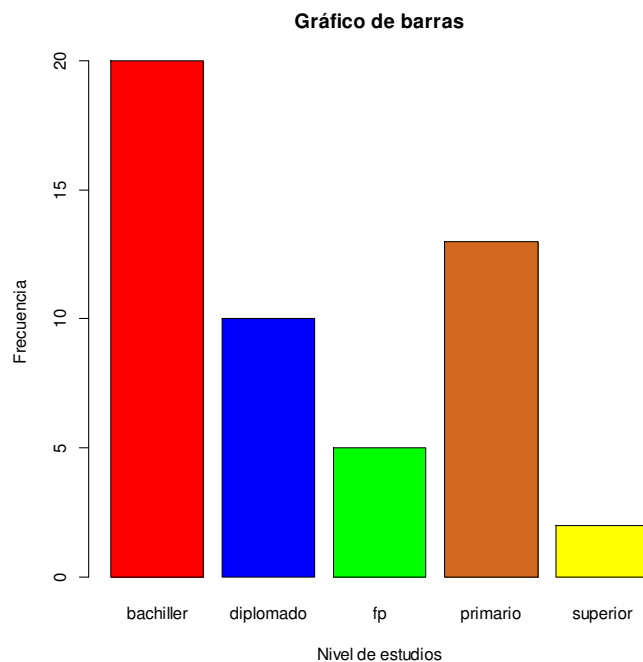
De nuevo, podemos modificar el código usado

```
barplot(table(Encuesta$estudios), xlab="estudios", ylab="Frequency")
```

para dar un nuevo aspecto al gráfico. Por ejemplo, cambiamos los nombres a los ejes, le damos un título al gráfico y le damos colores a cada barra. Así, ejecutando

```
barplot(table(Encuesta$estudios), xlab="Nivel de estudios", ylab="Frecuencia",
main="Gráfico de barras", col=c("red","blue","green","chocolate","yellow"))
```

el resultado es



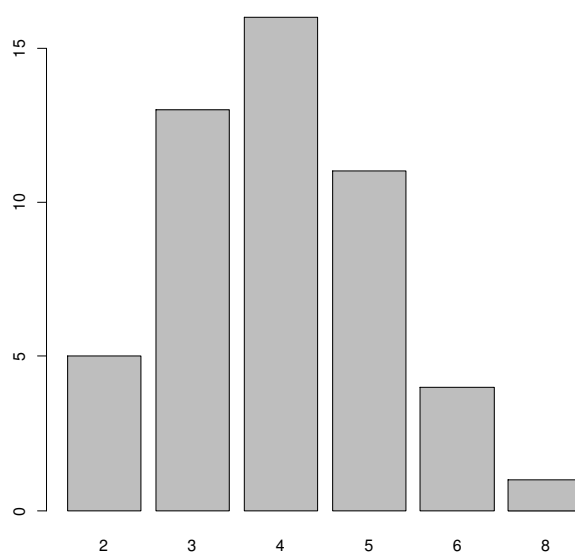
* Si queremos hacer un diagrama de barras para una variable discreta lo tendremos que hacer mediante código al igual que ocurría con la distribución de frecuencias.

Ejemplo

Para hacer un diagrama de barras con la variable “habitantes” del fichero *Encuesta*, ejecutamos la orden (sin ninguna opción)

```
barplot(table(Encuesta$habitantes))
```

que daría lugar a



A partir de la orden anterior podemos incluir todas las opciones que queramos para personalizar el gráfico.

Gráficas

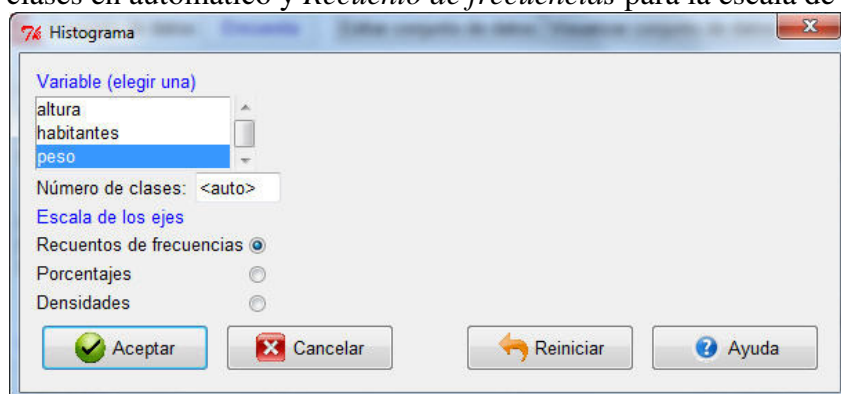
Histograma

Se selecciona una variable de tipo numérico (continua o discreta con un número elevado de valores), el número de clases y lo que queremos que aparezca contabilizado (altura de los rectángulos): recuento de frecuencias, porcentajes o densidades.

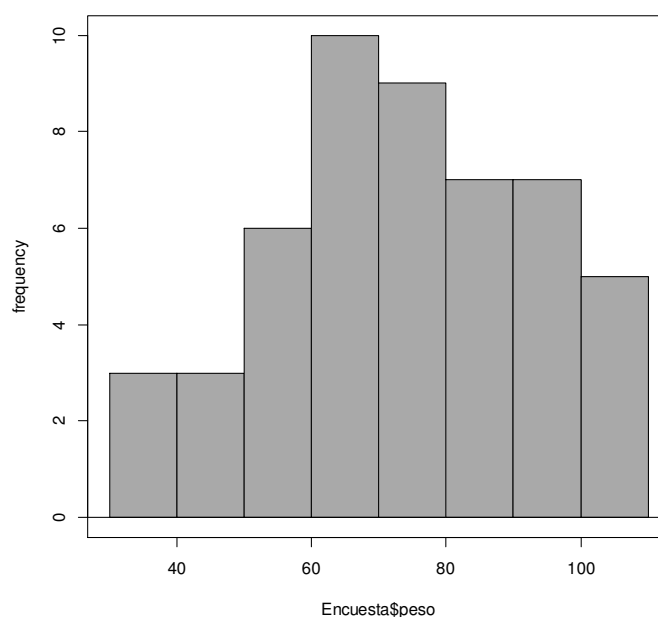
Al contabilizar las frecuencias de las clases toma los intervalos en la forma [], (), ..., ()

Ejemplo

Hacemos un histograma con la variable “peso” del fichero *Encuesta*. Con la opción de número de clases en automático y *Recuento de frecuencias* para la escala de los ejes



el resultado es



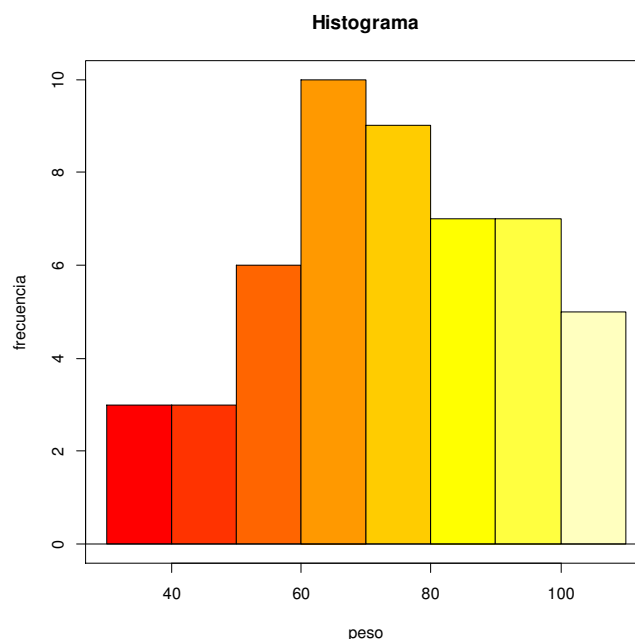
Si observamos la orden que se ha ejecutado

```
Hist(Encuesta$peso, scale="frequency", breaks="Sturges", col="darkgray")
```

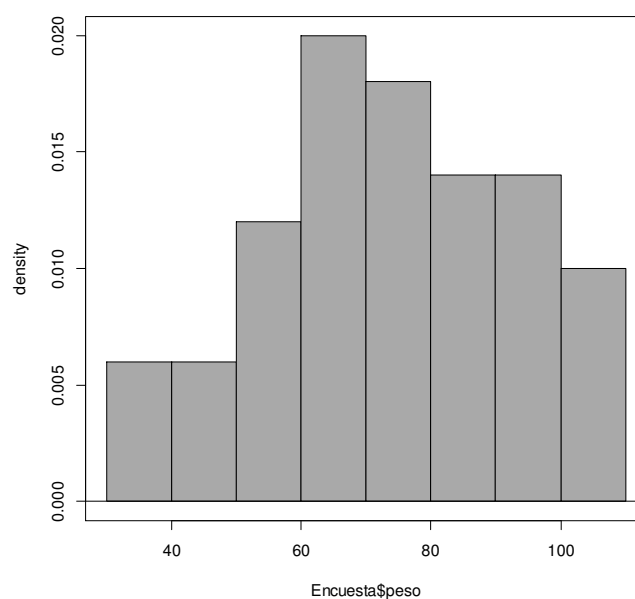
Podemos modificarla para personalizar el gráfico, por ejemplo

```
Hist(Encuesta$peso, scale="frequency", xlab="peso", ylab="frecuencia",
main="Histograma", breaks="Sturges", col= heat.colors(8))
```

da lugar a



Es interesante la realización de histogramas con la opción de *Densidades* para la escala de los ejes (de forma que la suma de todas las áreas de los rectángulos sea 1) para poder luego realizar comparaciones con funciones de densidad (funciones positivas que integran 1) asociadas con ciertas distribuciones de probabilidad.



* Podemos usar la siguiente orden (¡atención!, ahora hist va en minúscula)

```
hist(NombreFichero$NombreVariable, nclass = NumeroIntervalos, plot=FALSE)
```

para obtener una tabla de frecuencias para una variable numérica, agrupándola en clases.

Ejemplo

Al ejecutar

```
hist(Encuesta$peso, nclass = 8, plot=FALSE)
```

obtenemos una tabla de frecuencias de la variable “peso” (agrupada en 8 clases, las que ha dado por defecto el histograma). El resultado es

\$breaks

```
[1] 30 40 50 60 70 80 90 100 110
```

\$counts

```
[1] 3 3 6 10 9 7 7 5
```

\$intensities

```
[1] 0.006 0.006 0.012 0.020 0.018 0.014 0.014 0.010
```

\$density

```
[1] 0.006 0.006 0.012 0.020 0.018 0.014 0.014 0.010
```

\$mids

```
[1] 35 45 55 65 75 85 95 105
```

Gráficas

Gráfica de tallo y hojas

Se selecciona una variable de tipo numérico y las especificaciones adecuadas a los datos disponibles.

Ejemplo

Para obtener el siguiente gráfico para la variable “altura” del fichero *Encuesta*

n: 50

```
3  13 | 468
6  14 | 467
13 15 | 1223489
23 16 | 3455678899
(9) 17 | 024445688
18 18 | 000349
12 19 | 345668891010
2  20 | 02
```

se han seleccionado las siguientes opciones

