

Iniciación a R Commander*

John Fox

(21 de abril de 2008)

1. Iniciar R Commander

Una vez que R se está ejecutando, simplemente cargando el paquete `Rcmdr` mediante la instrucción `library(Rcmdr)` en la consola de R, se inicia la interfaz gráfica de usuario (“GUT”) de R Commander. Para que funcione correctamente con Windows, R Commander requiere la interfaz de documento único (SDI) de R¹. Después de cargar el paquete, las ventanas `R Console` y `R Commander` deberían aparecer más o menos como en las figuras 1 y 2. Éstas y otras imágenes en este documento fueron creadas con Windows XP; si usa otra versión de Windows (o, por supuesto, otra plataforma informática), entonces la apariencia de la pantalla puede diferir². Si estos paquetes no están instalados, `Rcmdr` ofrecerá instalarlos desde Internet o desde archivos locales (e.g., desde un CD-ROM). Si usted instala el paquete `Rcmdr` a través de Windows “R GUT”, los paquetes de los cuales `Rcmdr` depende deberían ser instalados automáticamente. En general, puede instalar el paquete `Rcmdr` y todos los paquetes de los que depende a través de la función `install.packages`, estableciendo el argumento `dependencies = TRUE`. Gracias a Dirk Eddelbuettel, los usuarios de Debian Linux sólo necesitan la instrucción `apt-get install r-cran-rcmdr` para instalar el paquete `Rcmdr` junto con todos los paquetes que éste requiere. En cualquier caso, instalar y cargar el paquete `Rcmdr` en los sistemas Linux suele ser sencillo. La tarea puede ser más complicada con OS/X en los sistemas Macintosh, ya que el paquete `telTk` del cual depende `Rcmdr` requiere que `Tcl/Tk` esté instalado y que R se esté ejecutando con X-Windows. Para habilitar los gráficos en 3D en `Rcmdr`, instale el paquete `rgl`. Puede encontrar más información sobre la instalación en la web de R Commander <http://socserv.socsci.mcmaster.ca/jfox/Misc/Rcmdr/index.html> y en http://knuth.uca.es/R/doku.php?id=r_wiki.

* Este manual está revisado y traducido en marzo de 2008, del original de Fox (2005), como parte del proyecto R-UCA, puede encontrar más información en <http://knuth.uca.es/R>.

¹La versión de Windows de R normalmente se ejecuta desde una interfaz de documentos múltiples (MDI), que contiene la ventana `R Console`, las ventanas `Graphics Device` creadas durante la sesión y cualquier otra ventana relacionada con el proceso R. En contraste, en la interfaz de documento único (“SDI”), las ventanas `R Console` y `Graphics Device` no están contenidas en una ventana maestra. Hay varias maneras de ejecutar R en modo SDI, por ejemplo, editando el archivo `Rconsole` en el subdirectorio `etc` de R, o añadiendo `-sdi` al campo `Destino` en la pestaña `Acceso Directo` dentro de `Propiedades` del icono del escritorio R. Esta limitación del paquete `Rcmdr` es heredada del paquete `telTk`, del cual depende `Rcmdr`.

²Dese cuenta de que `Rcmdr` requiere algunos paquetes además de varios de los paquetes “recomendados” que normalmente son distribuidos con R, y de que carga estos paquetes en el arranque. `Rcmdr`, los paquetes requeridos, y muchos otros paquetes aportados están disponibles para descarga desde la Red Extensa de Archivos R [Comprehensive R Archive Network (CRAN)] en <http://cran.r-project.org>

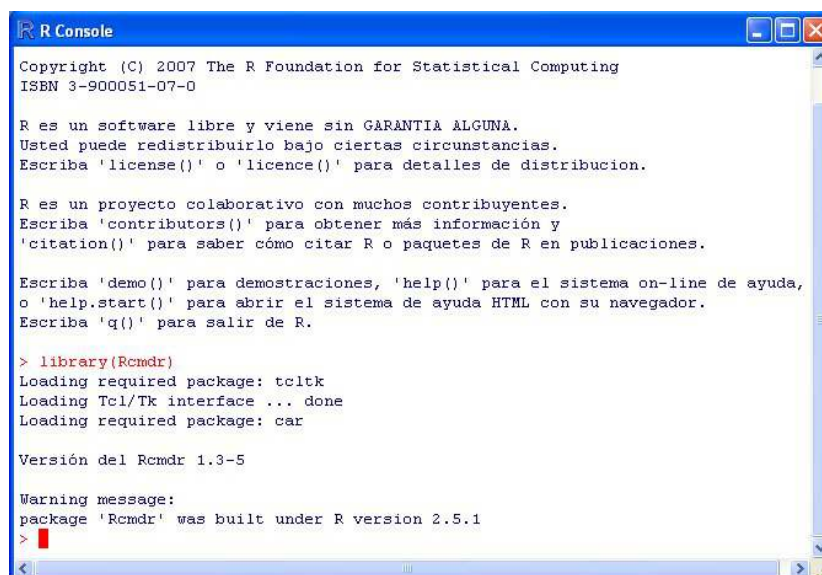


Figura 1: La ventana R Console tras cargar el paquete Rcmdr.

Las ventanas R Commander y R Console flotan libremente en el escritorio. Normalmente usted usará los menús y cuadros de diálogo de R Commander 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. También puede escribir las instrucciones R directamente en la ventana de instrucciones o en R Console tras el símbolo del sistema > (mayor que); sin embargo, el propósito principal de R Commander es evitar tener que escribir instrucciones.
- El resultado aparece impreso por defecto en la segunda ventana de texto (denominada Ventana de Resultados).
- La ventana gris inferior (denominada Mensajes) muestra mensajes de error, advertencias y otra información (notas), como el mensaje de inicio de la figura 2.
- Cuando cree gráficos, éstos aparecerán en una ventana separada Graphics Device.

Existen varios menús en la parte superior de la ventana R Commander:

Fichero Opciones de menú para cargar y guardar archivos de instrucciones; para guardar resultados y el área de trabajo R; y para salir.

Editar Opciones de menú (Cortar, Copiar, Pegar, etc.) para editar los contenidos de las ventanas de instrucciones y de resultados. Al pulsar con el botón derecho en la ventana de instrucciones o de resultados también aparece un menú “contextual” de edición.

Datos Submenús que contienen opciones de menú para leer y manipular datos.

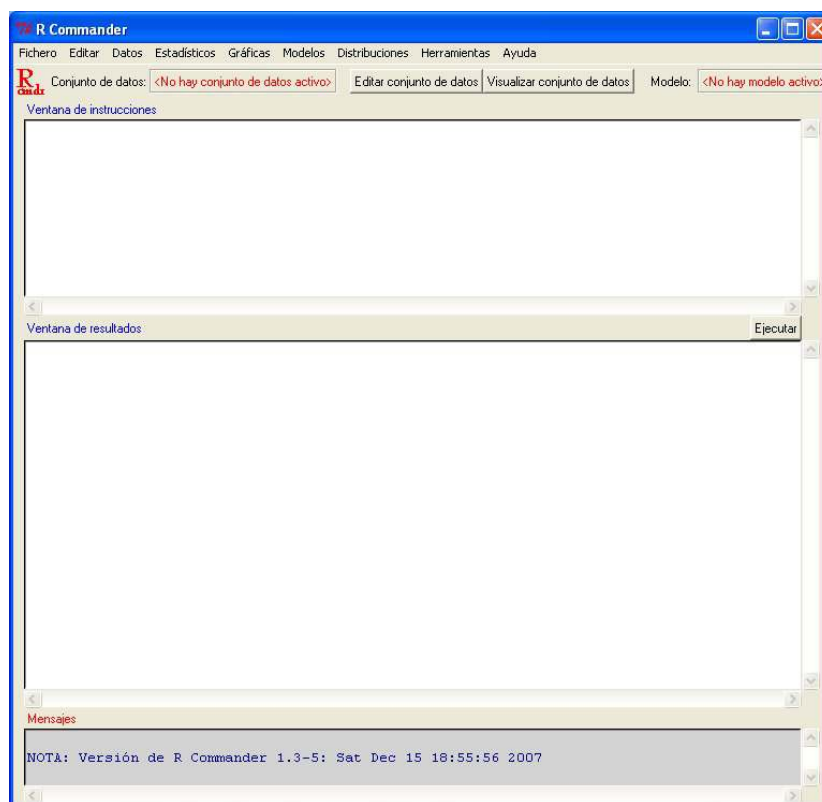


Figura 2: La ventana de R Commander al inicio.

Estadísticos Submenús que contienen opciones de menú para una variedad de análisis estadísticos básicos.

Gráficas Opciones de menú para crear gráficos estadísticos simples.

Modelos Opciones de menú y submenús para obtener resúmenes, intervalos de confianza, tests de hipótesis, diagnósticos y gráficas para un modelo estadístico, y para añadir cantidades diagnósticas, como residuos, a la serie de datos.

Distribuciones Probabilidades, cuantiles y gráficos para distribuciones estadísticas estándares (para usarse, por ejemplo, como sustituto de las tablas estadísticas) y ejemplos de estas distribuciones.

Herramientas Opciones de menú para cargar paquetes R no relacionados con el paquete Rcmdr (e.g., para acceder a datos guardados en otro paquete), y para establecer algunas opciones.

Ayuda Opciones de menú para obtener información sobre R Commander (incluyendo este manual). Además, cada cuadro de diálogo de R Commander tiene un botón de Ayuda (ver más abajo).

El menú “en árbol” completo de R Commander (versión 1.2-0) se muestra en el apéndice A. La mayoría de las opciones de menú conducen a cuadros de diálogo, como se muestra más adelante en este documento. Las opciones de menú están inactivas (se muestran en gris) si no son aplicables en el contexto actual.

La interfaz de R Commander incluye algunos elementos además de los menús

y diálogos:

- Bajo los menús hay una “barra de herramientas” con una fila de botones.
 - El botón horizontal que se encuentra más a la izquierda muestra el nombre de la serie de datos activa. Inicialmente no hay ninguna serie de datos activa. Si presiona este botón, podrá elegir entre las series de datos que están actualmente en la memoria (si hay más de una). La mayoría de los menús y diálogos en R Commander se refieren a la serie de datos activa (los menús Fichero, Editar y Distribuciones son excepciones).
 - Los dos botones contiguos le permiten abrir el editor de datos de R para modificar la serie de datos activa o el visor para examinarla. El visor de la serie de datos puede permanecer abierto mientras se realizan otras operaciones³.
 - Un botón horizontal indica el nombre del modelo estadístico activo, un modelo lineal (como el modelo de regresión lineal), un modelo lineal generalizado, un modelo logit multinominal o un modelo lineal de regresión⁴. Inicialmente no hay ningún modelo activo. Si hay más de un modelo en la memoria, puede elegir entre ellos pulsando sobre el botón.
- Inmediatamente debajo de la barra de herramientas se encuentra la ventana de instrucciones (denominada de esta manera), una ventana grande en la cual podrá desplazarse hacia arriba y abajo en el texto. Como se ha mencionado anteriormente, las instrucciones generadas por la GUI se copian en esta ventana. Puede editar el texto en la ventana de instrucciones o incluso escribir sus propias instrucciones R en la ventana. Presionando el botón Ejecutar, que está a la derecha debajo de la ventana de instrucciones (o, alternativamente, la combinación de teclas Ctrl-r para “ejecutar”), la línea contenida en el cursor se ejecuta (o se re-ejecuta) para su resolución. Si se seleccionan varias líneas (e.g., pulsando con el botón izquierdo y arrastrando el ratón sobre ellas), el presionar Ejecutar hará que todas ellas se ejecuten. Las instrucciones introducidas en la ventana de instrucciones pueden ocupar más de una línea, pero si lo hacen, las líneas siguientes a la primera deberán sangrarse con uno o más espacios o tabulaciones. La combinación de teclas Ctrl-a selecciona todo el texto de la ventana de instrucciones, y Ctrl-s abre una ventana de diálogo para guardar los contenidos de la ventana.
- Bajo la ventana de instrucciones hay una ventana grande para los resultados que permite editar el texto y desplazarse arriba y abajo sobre él. Las

³El visor de datos, proporcionado por la función `showData` del paquete `relimp` de David Firth, puede ir lento con series de datos con gran número de variables. Cuando el número de variables excede el límite (inicialmente fijado en 100), se usa el editor de datos R en su lugar para mostrar la serie de datos. Para usar el editor de datos sin tener en cuenta el número de variables, fije el límite en 0. Vea la ayuda de R Commander para más detalles. Una desventaja de usar el editor de datos para mostrar la serie de datos actual es que la ventana de edición no se puede seguir mostrando mientras se realizan otras operaciones.

⁴Los usuarios pueden proporcionar clases adicionales de modelos estadísticos añadiendo los cuadros de diálogo necesarios y las opciones de menú, y editando el archivo `model-classes.txt` en el directorio `etc` de R.

instrucciones escritas en esta ventana aparecen en rojo y los resultados en azul oscuro (como en la R Console).

- Al final hay una pequeña ventana de texto gris para mensajes. Los mensajes de error se muestran en rojo, los avisos en verde y otros mensajes en azul oscuro.

Una vez que ha cargado el paquete Rcmdr, puede minimizar la R Console. La ventana R Commander puede ser reajustada o maximizada de la manera usual. Si reajusta R Commander, el ancho de los resultados R subsiguientes automáticamente se ajustará para encajar en la ventana de resultados. R Commander es altamente configurable, la configuración por defecto se ha descrito aquí. Se pueden realizar cambios en esta configuración a través del menú **Herramientas**→**Opciones**, o, mucho más extensamente, estableciendo las opciones en R.⁵ Ver la ayuda de Rcmdr para más detalles.

2. Entrada de datos

La mayoría de los procesos en R Commander asumen que hay un conjunto de datos activo⁶. Si hay varios conjuntos de datos en la memoria, podrá elegir entre ellos, pero sólo uno estará activo. Cuando R Commander se inicia, no hay ningún conjunto de datos activo. R Commander proporciona varias maneras de introducir datos en R:

- Puede introducir los datos directamente a través de **Datos**→**Nuevo** conjunto de datos. . . Esta es una opción razonable para un conjunto de datos pequeño.
- Puede importar datos desde un archivo de texto sin formato (“ascii”) o el portapapeles, desde otro paquete estadístico (Minitab, SPSS o Stats), o desde un conjunto de datos de Excel, Access o dBase.
- Puede leer un conjunto de datos que esté incluido en un paquete R, ya sea escribiendo el nombre del conjunto de datos (si lo conoce) o seleccionando el conjunto de datos en un cuadro de diálogo.

2.1. Leer datos desde un archivo de texto

Por ejemplo, considere el archivo de datos Nations.txt⁷ Las primeras líneas del archivo son las siguientes:

	TFR	contraception	infant.mortality	GDP	region
Afghanistan	6.90	NA	154	2848	Asia
Albania	2.60	NA	32	863	Europe
Algeria	3.81	52	44	1531	Africa

⁵Una opción de menú que termine en elipses (p.e., tres puntos, . . .) conduce a un cuadro de diálogo; esto es un consenso estándar de la GUI. En este documento→representa una opción de menú o un submenú de un menú.

⁶Los procesos seleccionados mediante el menú Distribuciones son excepciones, así como Introducir y analizar tablas de doble entrada a través del menú Estadísticos→Tablas de contingencia.

⁷Este archivo se encuentra en el subdirectorio etc del paquete Rcmdr.

American-Samoa	NA	NA	11	NA Oceania
Andorra	NA	NA	NA	NA Europe
Angola	6.69	NA	124	355 Africa
Antigua	NA	53	24	6966 Americas
Argentina	2.62	NA	22	8055 Americas
Armenia	1.70	22	25	354 Europe
Australia	1.89	76	6	20046 Oceania

- La primera línea del archivo contiene los nombres de las variables: TFR (tasa total de fertilidad, expresada como el número total de niños por mujer), contraception (la tasa de uso de contraceptivos entre las mujeres casadas, en porcentaje), infant.mortality (la tasa de mortalidad infantil por cada 1000 nacidos vivos), GDP (producto doméstico bruto per cápita, en dólares americanos) y region (región).
- Las líneas siguientes contienen los valores de estos datos, una línea por país. Los valores de datos están separados por “espacio en blanco”, uno o más espacios o tabulaciones. Aunque es útil alinear los datos verticalmente, no es necesario hacerlo. Dese cuenta de que las líneas de datos comienzan con los nombres de los países. Como queremos que éstos sean los “nombres de cada fila” para el conjunto de datos, no hay un nombre de variable que le corresponda: Esto es, hay cinco nombres de variables pero seis valores en cada línea. Cuando esto sucede, R interpreta que el primer valor de cada línea es el nombre de la fila.
- Algunos de los valores faltan. En R, lo más conveniente es usar NA (que representa “no disponible” (not available) para codificar los valores que faltan, como se ha hecho aquí.
- Las variables TFR, contraception, infant.mortality y GDP son variables numéricas (cuantitativas); por el contrario, region contiene nombres de regiones. Cuando los datos son leídos, R tratará region como un “factor”, es decir, una variable categórica. En la mayoría de los contextos, R Commander distingue entre las variables numéricas y los factores.

Para leer un archivo de datos en R, seleccionar **Datos** → **Importar datos** → **desde archivo de texto o portapapeles...** Esta operación abre un cuadro de diálogo llamado Leer datos desde archivo de texto o portapapeles, como se muestra en la figura 3. Los nombres R válidos comienzan con una letra mayúscula o minúscula (o un punto, .) y consisten enteramente en letras, puntos, subrayados (_) y números (p.e., 0-9); en particular, no se incluirá ningún espacio en un nombre de conjunto de datos. También debería saber que R es sensible a las mayúsculas, por lo que, por ejemplo, se distingue entre naciones, Naciones y NACIONES, y podría usarse para representar diferentes conjuntos de datos. Pulsando sobre el botón Aceptar que se encuentra en Leer datos desde archivo de texto o portapapeles se abre el cuadro de diálogo Abrir para abrir archivos, como se muestra en la figura 4. Nos dirigimos hacia el archivo Nations.txt. Presionando el botón Abrir se leerá el archivo de datos. Una vez que éste se ha leído, se convierte en el conjunto de datos activo en R Commander. En consecuencia (figura 5), el nombre del conjunto de datos aparece en el botón de conjunto de datos que está arriba a la izquierda en la ventana R Commander. Pulsamos en

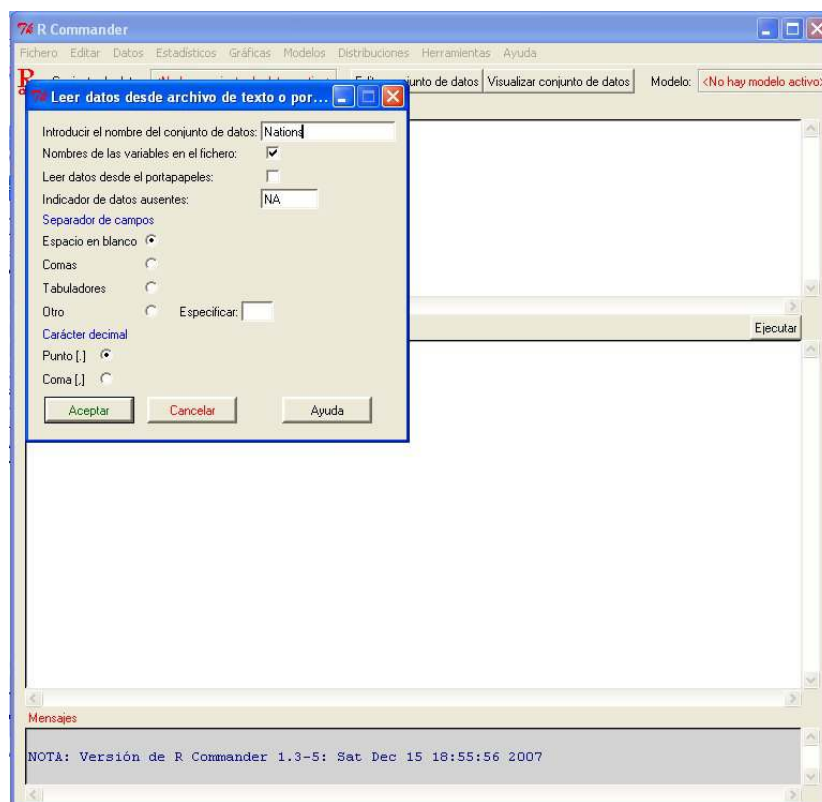


Figura 3: Lectura de datos desde un archivo de texto.

el botón Visualizar conjunto de datos para abrir la ventana del visor de datos, como también se muestra en la figura 5. Dese cuenta de que las instrucciones para leer y ver el conjunto de datos Nations (las instrucciones de R `read.table` y `showData`) aparecen, parcialmente tapados por la visualización del conjunto de datos, en las ventanas de instrucciones y de resultados. Cuando el conjunto de datos se lee y se convierte en el conjunto de datos activo, una nota aparece en la ventana de mensajes (que desaparece cuando la siguiente instrucción `showData` se ejecuta).

La instrucción `read.table` crea un “conjunto de datos” R, que es un objeto que contiene un conjunto de datos en el formato casos por variables: Las filas del conjunto de datos representan los casos o las observaciones y las columnas las variables. Los conjuntos de datos en R Commander son conjuntos de datos en R.

2.2. Introducir datos directamente

Para introducir datos directamente en el editor de datos, semejante a una hoja de cálculo, puede proceder como se explica a continuación. Para ilustrar ésto, hemos usado un conjunto de datos muy pequeño del problema 2.44 de Moore (2000):

- Seleccione Datos→Nuevo conjunto de datos... en los menús de R Com-

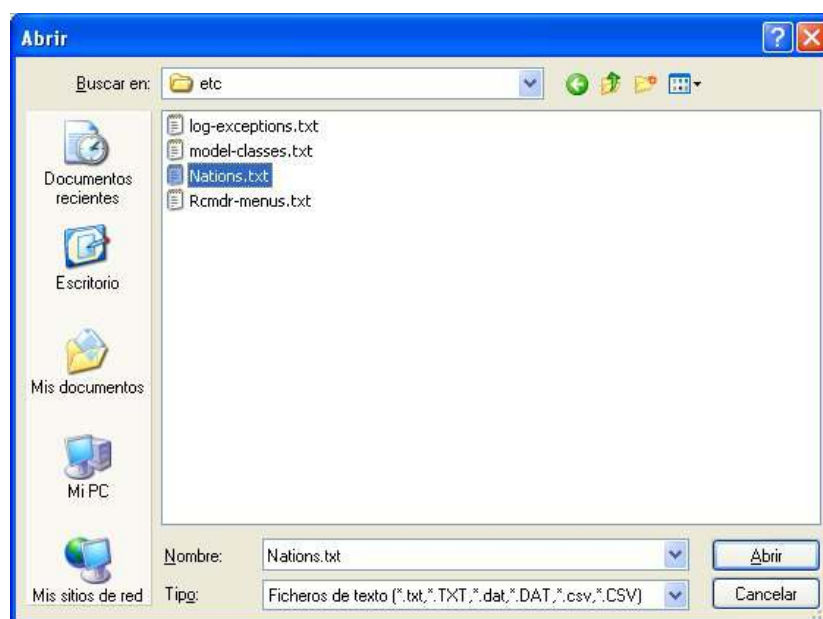


Figura 4: Diálogo para abrir un archivo para la lectura de datos en un archivo de texto.

mander. Opcionalmente puede introducir un nombre para el conjunto de datos, como por ejemplo Problem2.44, en el cuadro de diálogo resultante, y luego pulsar en el botón Aceptar. (Recuerde que los nombres en R no pueden incluir espacios). Esto abrirá la ventana Editor de datos con un conjunto de datos vacío.

- Introduzca los datos del problema en las dos primeras columnas del editor de datos. Puede desplazarse de una celda a otra usando las flechas del teclado, presionando Tab o Enter, o situándose sobre la celda con el ratón y pulsando con el botón izquierdo. Cuando haya acabado de introducir los datos, la ventana debería parecerse a la figura 6.
- A continuación, pulse sobre el nombre var1 en la primera columna. Esto abrirá un cuadro de diálogo llamado Editor de variables, como en la figura 7.
- Escriba como nombre de la variable edad, justo como hemos hecho, y pulse sobre la X que se encuentra en la esquina superior derecha de la ventana Editor de variables, o presione la tecla Enter, para cerrar la ventana. Repita este procedimiento para nominar a la segunda columna height. El Editor de datos debería parecerse ahora a la figura 8.
- Seleccione **Fichero**→**Cerrar** en los menús del Editor de datos o pulse en la X de la esquina superior derecha de la ventana Editor de datos. El conjunto de datos que ha introducido es ahora el conjunto de datos activo en R Commander.

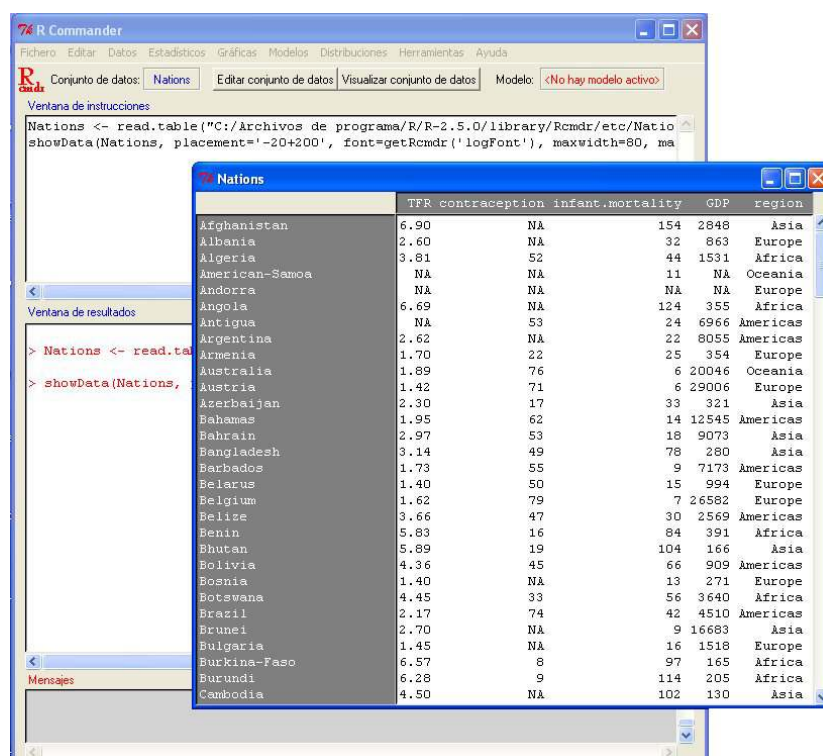


Figura 5: Mostrando el conjunto de datos activo.

2.3. Lectura de datos desde un paquete

Muchos paquetes R contienen datos. Los conjuntos de datos en paquetes pueden listarse en una ventana desplegable a través de **Datos**→**Conjunto de datos en paquetes**→**Leer conjunto de datos desde paquete adjunto**.⁸ El cuadro de diálogo resultante se muestra en la figura 9. Si sabe el nombre de un conjunto de datos en un paquete, entonces puede introducir el nombre directamente; en caso contrario, al hacer una doble pulsación sobre el nombre de un paquete se muestra su conjunto de datos listado en el cuadro de la derecha; y haciendo una doble pulsación sobre el nombre del conjunto de datos se copia el nombre al campo de entrada del conjunto de datos.⁹ Puede adjuntar paquetes R adicionales en **Herramientas**→**Cargar paquete(s)**.

3. Crear resúmenes numéricos y gráficas

Una vez que hay un conjunto de datos activo, puede usar los menús de R Commander para generar varios resúmenes numéricos y gráficas. Describiremos simplemente unos ejemplos básicos a continuación. Una buena GUI debería

⁸No todos los datos en paquetes son estructuras de datos y sólo las estructuras de datos son adecuadas para usarse en R Commander. Si intenta leer datos que no se encuentran en estructuras de datos, aparecerá un mensaje de error en la ventana de mensajes de abajo.

⁹En general en R Commander, cuando es necesario copiar un dato desde una lista a otra localización en el diálogo, basta con una doble pulsación.

	var1	var2	var3	var4	var5	var6
1	36	86				
2	48	90				
3	51	91				
4	54	93				
5	57	94				
6	60	96				
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						

Figura 6: Editor de datos una vez que se han introducido los datos.

Figura 7: Cuadro de diálogo para cambiar el nombre de una variable en el editor de datos.

ser ampliamente autoexplicativa: esperamos que una vez que vea cómo funciona R Commander, le sea fácil usarlo, asistido, quizás, por la ayuda en línea. En los ejemplos de más abajo, asumimos que el conjunto de datos activo es Nations, leído desde un archivo de texto en la sección anterior. Si por el contrario escribió los datos manualmente desde Moore (2000) siguiendo los cinco pasos descritos arriba, o leyó el conjunto de datos Prestige del paquete car, como también se describió más arriba, entonces uno de estos es el conjunto de datos activo. Recuerde que puede cambiar el conjunto de datos activo pulsando en el botón horizontal con el nombre del conjunto de datos activo que se encuentra arriba a la izquierda de la ventana R Commander, seleccionando entre los conjuntos de datos listados actualmente en la memoria. Seleccionando **Estadísticos**→**Resúmenes**→**Conjunto de datos activo**, obtendrá la figura 10. Para cada variable numérica en el conjunto de datos (TFR, contraception, infant.mortality y GDP), R informa de los valores máximos y mínimos, el primer y el tercer cuartil, la mediana, la media, así como el número de valores que faltan. Para la variable categórica region, obtenemos el número de observaciones en cada “nivel” del factor. Si el conjunto de datos hubiera incluido más de diez variables, R Commander nos habría preguntado si realmente queremos continuar, intentando protegernos de producir una no deseada voluminosa cantidad de resultados. De manera similar, seleccionando **Estadísticos**→**Resúmenes**→**Resúmenes numéricos...** se abre el cuadro de diálogo mostrado en la figura 11. Sólo las variables numéricas se muestran en la lista de variables de este diálogo; el factor región falta, porque no es razonable calcular resúmenes numéricos para un factor. Pulsando en infant.mortality y a continuación en Aceptar se obtiene el

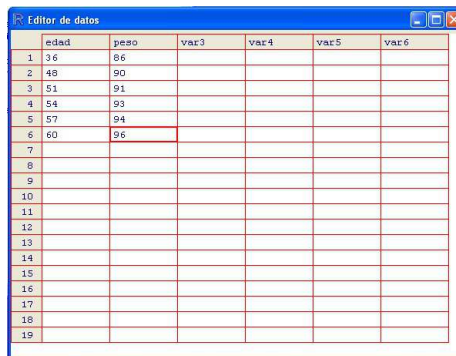


Figura 8: La ventana Editor de datos tras haber cambiado los nombres de las dos variables.



Figura 9: Lectura de datos desde un paquete adjunto.

siguiente resultado (en la ventana de resultados).¹⁰

```
> numSummary(Nations[,"infant.mortality"],
statistics=c("mean", "sd", "quantiles"))
mean sd 0 % 25 % 50 % 75 % 100 % n NA
43.47761 38.75604 2 12 30 66 169 201 6
```

Por defecto, la instrucción R que se ejecuta calcula la media y la desviación estándar (sd) de la variable, junto con los cuantiles correspondientes con el mínimo, el primer cuartil, la mediana, el tercer cuartil y el máximo; n es el número de observaciones válidas, y NA es el número de valores que faltan. Como es típico en los diálogos de R Commander, el cuadro de diálogo resúmenes numéricos de la figura 11 incluye los botones Aceptar, Cancelar y Ayuda. El botón Ayuda conduce a una página de ayuda, ya sea para el diálogo en sí o (como en este caso) para la función R a la que el diálogo alude. El cuadro de diálogo resúmenes numéricos también tiene en cuenta para calcular resúmenes los grupos definidos por los niveles de un factor. Pulsando en el botón Resumir por grupos... se abre el diálogo Grupos, como se muestra en la figura 12. Debido a que sólo hay un factor en el conjunto de datos Naciones, solamente la variable

¹⁰Para seleccionar una sola variable en la lista de variables del cuadro, simplemente pulse con el botón izquierdo sobre su nombre. En algunos contextos, tendrá que (o querrá) seleccionar más de una variable. En estos casos, se aplican los consensos de Windows habituales: pulsando con el botón izquierdo del ratón sobre una variable la selecciona y también deshace la selección de cualquier variable que haya sido previamente seleccionada; presionando la tecla mayúsculas mientras pulsa con el botón izquierdo extenderá la selección; y Ctrl + botón izquierdo conmuta la selección por una variable individual.

región aparecerá en la lista de variables; al seleccionar esta variable y pulsando en Aceptar se cambia el botón Resumir por grupos a Resumir por region (ver figura 13); al pulsar en Aceptar aparece el siguiente resultado:

```
> numSummary(Nations[, "infant.mortality"],
             groups=Nations$region, statistics=c("mean", "sd", "quantiles"))
```

	mean	sd	0 %	25 %	50 %	75 %	100 %	n	NA
Africa	85.27273	35.188095	7	61.00	85.0	111.00	169	55	0
Americas	25.60000	17.439713	6	12.00	21.5	36.00	82	40	1
Asia	45.65854	32.980001	5	22.00	37.0	72.00	154	41	0
Europe	11.85366	7.122363	5	6.00	8.0	16.00	32	41	4
Oceania	27.79167	29.622229	2	9.25	20.0	35.75	135	24	1

Otros diálogos de R Commander le permiten seleccionar una variable grupo de esta manera.

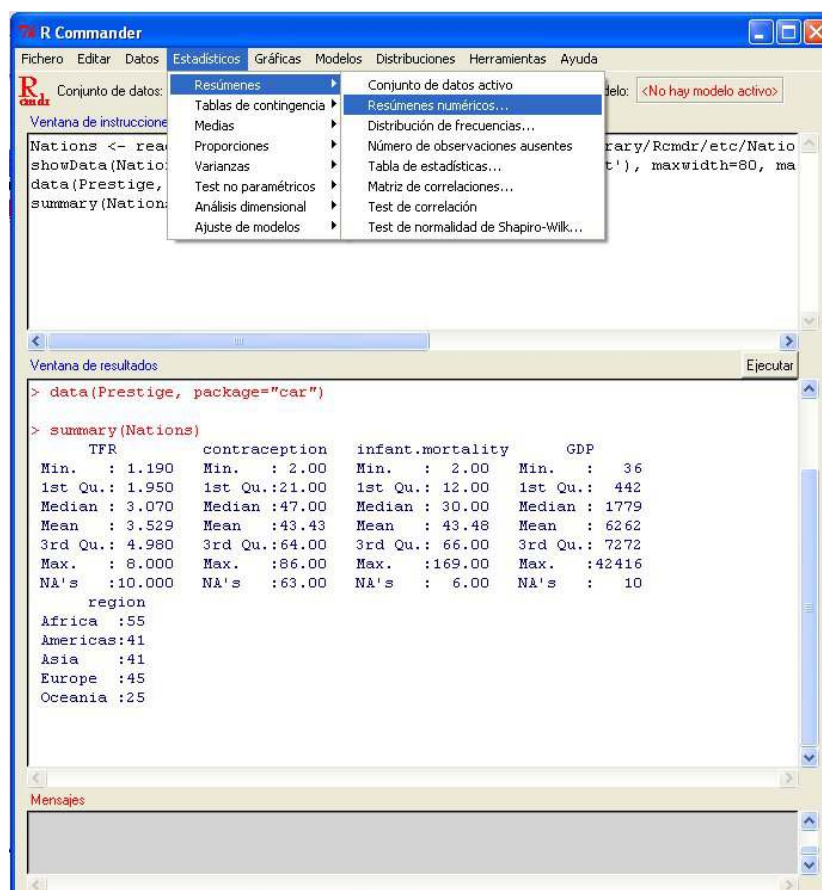


Figura 10: Obtener resúmenes de variables para el conjunto de datos activo.

Hacer gráficas con R Commander también es sencillo. Por ejemplo, seleccionando Gráficas→Histograma... en los menús de R Commander se abre el cuadro de diálogo Histograma de la figura 14; y pulsando en infant.mortality



Figura 11: Cuadro de diálogo Resúmenes Numéricos.



Figura 12: Selección de una variable agrupadora en el cuadro de diálogo Grupos.

seguido de Aceptar se abre la ventana Graphics Device con el histograma que se muestra en la figura 15. Si hace varias gráficas en una sesión, normalmente sólo aparecerá la más reciente en la ventana Graphics Device. Puede recuperar las gráficas previas usando las teclas Page Up y Page Down de su teclado¹¹.

4. Modelos estadísticos

Varios tipos de modelos Estadísticos pueden ajustarse en R Commander usando las opciones de menú **Estadísticos**→**Ajuste de modelos**: modelos lineales (a través de Regresión Lineal y Modelo lineal), modelos lineales generalizados, modelos logit multinomial y modelos ordinales de regresión [los dos últimos desde los paquetes *nnet* y *MASS* de Venables y Ripley (2002), respectivamente]. Aunque los cuadros de diálogo resultantes difieren en ciertos detalles (por ejemplo, el diálogo del modelo lineal generalizado facilita la selección de

¹¹Al inicio, R Commander pone en marcha el mecanismo del historial gráfico; esta característica sólo está disponible en los sistemas con Windows. Los diagramas de dispersión dinámicos tridimensionales creados mediante Gráficas → Gráficas 3D → Diagrama de dispersión en 3D... aparecen en la ventana especial RGL device; de la misma manera, las muestras de efecto creadas para modelos estadísticos (Fox, 2003) a través de Modelos → Gráficas → Gráfica de los efectos... aparecen en una ventana individual de dispositivo gráfico.



Figura 13: Cuadro de diálogo Resúmenes Numéricos tras seleccionar una variable agrupadora.

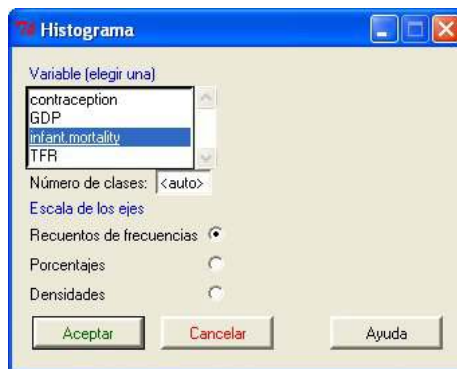


Figura 14: Diálogo Histograma.

una familia de distribución y el correspondiente link para la función), comparten una estructura general común, como se ilustra en el diálogo Modelo lineal de la figura 16.¹²

- Al hacer una doble pulsación sobre una variable en el cuadro con la lista de variables la copia a la fórmula del modelo, al lado izquierdo de la fórmula, si está vacío; en caso contrario, al lado derecho (con un signo + precediéndolo si el contexto lo requiere). Dese cuenta de que los factores (variables categóricas) aparecen etiquetados como tales entre paréntesis en la lista de variables.
- La fila de botones sobre la fórmula puede usarse para introducir operadores y paréntesis en el lado derecho de la fórmula.
- También puede teclear directamente en los campos para las fórmulas y, de hecho, tiene que hacerlo para, por ejemplo, introducir un término tal como `log(income)` en la fórmula.

¹²Una excepción es el diálogo Regresión lineal, en el cual la variable de respuesta y las variables explicativas se seleccionan simplemente por el nombre desde una lista que contiene las variables numéricas del conjunto de datos actual. La explicación siguiente asume la familiaridad con la fórmula del modelo R; ver, por ejemplo, el manual Introducción a R que viene con R, al cual puede acceder a través del menú Ayuda en R Console.

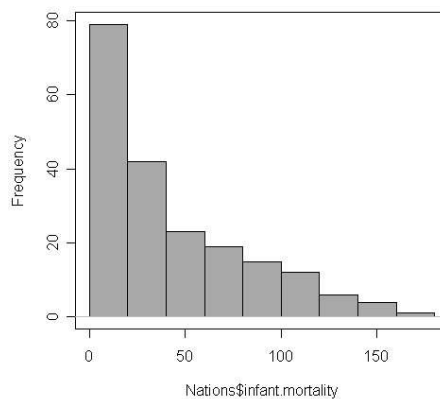


Figura 15: Una ventana de gráficas que contiene el histograma para mortalidad infantil.



Figura 16: Cuadro de diálogo Modelo lineal.

- El nombre del modelo, aquí LinearModel.1, es generado automáticamente, pero puede sustituirlo por cualquier nombre R válido.
- Puede escribir una expresión R en el cuadro llamado Expresión de selección; si está permitido, pasa al argumento subset de la función lm, y se usa para ajustar el modelo a un subconjunto de las observaciones en el conjunto de datos. Una forma de expresión del subconjunto es la expresión lógica que evalúa cada observación como TRUE o FALSE, como teclear != "prof" (lo cual seleccionaría todas las ocupaciones no profesionales del conjunto de datos Prestige).

Pulsando el botón Aceptar aparece el siguiente resultado (en la ventana de resultados), y hace que LinearModel.1 sea el modelo activo, con su nombre en el botón Modelo:

```
> LinearModel.1 <- lm(prestige~ (education + income)*type , data=Prestige)
> summary(LinearModel.1)
```

Call:
lm(formula = prestige~(education + income) * type, data = Prestige)

Residuals:
Min 1Q Median 3Q Max
-13.462 -4.225 1.346 3.826 19.631

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.276e+00	7.057e+00	0.323	0.7478
education	1.713e+00	9.572e-01	1.790	0.0769 .
income	3.522e-03	5.563e-04	6.332	9.62e-09 ***
type[T.prof]	1.535e+01	1.372e+01	1.119	0.2660
type[T.wc]	-3.354e+01	1.765e+01	-1.900	0.0607 .
education:type[T.prof]	1.388e+00	1.289e+00	1.077	0.2844
education:type[T.wc]	4.291e+00	1.757e+00	2.442	0.0166 *
income:type[T.prof]	-2.903e-03	5.989e-04	-4.847	5.28e-06 ***
income:type[T.wc]	-2.072e-03	8.940e-04	-2.318	0.0228 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 6.318 on 89 degrees of freedom
Multiple R-Squared: 0.8747, Adjusted R-squared: 0.8634
F-statistic: 77.64 on 8 and 89 DF, p-value: > 2.2e-16

Las operaciones en el modelo activo pueden seleccionarse en el menú Modelos. Por ejemplo, Modelos→Test de hipótesis→Tabla ANOVA da el siguiente resultado:

```
> Anova(LinearModel.1)
Anova Table (Type II tests)
Response: prestige
```

	Sum Sq	Df	F value	Pr(>F)
education	1068.0	1	26.7532	1.413e-06 ***
income	1131.9	1	28.3544	7.511e-07 ***
type	591.2	2	7.4044	0.001060 **
education:type	238.4	2	2.9859	0.055574 .
income:type	951.8	2	11.9210	2.588e-05 ***
Residuals	3552.9	89		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5. Excepciones y final

5.1. Salvar y guardar resultados

Puede guardar sus resultados de texto usando el menú **Fichero** en **R Commander**; de la misma manera, puede guardar o imprimir un gráfico desde el menú **Fichero** de la ventana **Graphics Device**. Generalmente es más conveniente, sin embargo, llevar el resultado de texto y los gráficos que quiera guardar a un documento de un procesador de texto. De esta manera, puede intercalar los resultados de R con sus notas y explicaciones propias. Abra un procesador de texto, como **Word**, o incluso **Windows WordPad**. Para copiar el texto de la ventana de resultados, seleccione el texto con el ratón, y elija **Copiar** del menú **Editar** (o presione la combinación de teclas **Ctrl-c**, o pulse con el botón derecho en la ventana y seleccione **Copiar** en el menú contextual), y luego pegue el texto en la ventana del procesador de texto a través de **Edición**→**Pegar** (o **Ctrl-v**), tal y como haría para cualquier aplicación de **Windows**. Un punto que merece la pena mencionar es que también debería usar una fuente monoespaciada (“typewriter”), como **Courier New**, para los resultados de R; de otra manera el resultado no se alinearán adecuadamente. De la misma manera, para copiar un gráfico, seleccione **Fichero**→**Copiar al portapapeles**→como **Metarchivo** en los menús de **Graphics Device**; luego pegue el gráfico en su procesador de texto mediante **Editar**→**Pegar** (o **Ctrl-v**). Alternativamente, puede usar **Ctrl-w** para copiar el gráfico desde **Graphics Device** de R, o pulsar con el botón derecho sobre el gráfico para abrir un menú contextual, en el cual podrá seleccionar **Copiar como metarchivo**¹³. Al final de su sesión en R puede guardar o imprimir el documento que ha creado, obteniendo un registro anotado de su trabajo. Puede encontrar otras rutas alternativas para guardar resultados de texto y gráficos en los menús de **R Commander Fichero** y **Gráficos**→**Guardar gráfico en archivo**, respectivamente.

5.2. Terminar la sesión en R

Hay varias maneras de terminar su sesión en R. Por ejemplo, puede seleccionar **Fichero**→**Salir**→**de Commander y R** en los menús de **R Commander**. Se le pedirá que confirme si desea salir, y a continuación si quiere guardar los contenidos de la ventana de instrucciones y de resultados. Del mismo modo, puede seleccionar **Fichero**→**Salir** en **R Console**; en este caso, se le preguntará si quiere guardar el área de trabajo R (p.e., los datos que R guarda en la memoria); normalmente contestará **No**.

5.3. Introducir instrucciones en la ventana de instrucciones

La ventana de instrucciones proporciona facilidades sencillas para editar, introducir y ejecutar instrucciones. Las instrucciones generadas por **R Commander** aparecen en la ventana de instrucciones, y usted puede teclear y editar instrucciones en la ventana más o menos como en cualquier editor. Sin embargo,

¹³Como podrá ver cuando examine estos menús, puede guardar los gráficos en una variedad de formatos

R Commander no proporciona una verdadera “consola” para R, y la ventana de instrucciones tiene algunas limitaciones:

- Las instrucciones que se extiendan más allá de una línea deberían tener la segunda y siguientes líneas sangradas con uno o más espacios o tabulaciones; todas las líneas de una instrucción multilínea deberán ser ejecutadas simultáneamente para obtener resultados.
- Las instrucciones que incluyan una flecha de asignación ($< -$) no generarán un resultado impreso, incluso si este resultado habría aparecido normalmente si se hubiera introducido la instrucción en R Console [la instrucción `print(x < -10)`, por ejemplo]. Por otro lado, las tareas hechas con el signo igual ($=$) producen una salida impresa incluso cuando normalmente no lo harían (e.g., `x = 10`).
- Las instrucciones que normalmente dan resultados invisibles ocasionalmente harán que los resultados aparezcan en la ventana de resultados. Este comportamiento puede modificarse editando las entradas del archivo `log-exceptions.txt` en el directorio `etc` de R Commander.
- Los bloques de instrucciones entre paréntesis, p.e., `(...)`, no son tratados adecuadamente a no ser que cada instrucción termine con un punto y coma (;). Este es el estilo R pobre, y conlleva que la ventana de instrucciones tenga un uso limitado como editor de programas. Para programación R seria, sería preferible usar el editor de instrucciones proporcionado por la versión de R para Windows, o, aún mejor, un editor de programas.

Referencias

- [1] Fox, J. (2003). Effect displays in R for generalised linear models. *Journal of Statistical Software*, 8(15):1-27.
- [2] Fox, J. (2005). The R Commander: A basic-statistics graphical user interface to R. *Journal of Statistical Software*, 19(9):1-42.
- [3] Moore, D. S. (2000). *The Basic Practice of Statistics*, Second Edition. Freeman, New York.
- [4] Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S*, Fourth Edition. Springer, New York.

A. Menú Rcmdr

```
Fichero - Abrir fichero de instrucciones...
| - Guardar las instrucciones
| - Guardar las instrucciones como...
| - Guardar los resultados...
| - Guardar los resultados como...
| - Guardar el entorno de trabajo R...
| - Guardar el entorno de trabajo R como...
| - Salir - De Commander
```

| - De Commander y R

Editar - Limpiar ventana

- | - Cortar
- | - Copiar
- | - Pegar
- | - Borrar
- | - Buscar...
- | - Seleccionar todo

Datos - Nuevo conjunto de datos...

- | - Cargar conjunto de datos...
- | - Importar datos - desde archivo de texto o portapapeles...
 - | - desde datos SPSS...
 - | - desde datos Minitab
 - | - desde datos STATA...
 - | - desde Excel, Access o dBase...
- | Conjunto de datos en paquetes - Lista de conjuntos de datos en paquetes
 - | - Leer conjunto de datos desde paquete adjunto...
- | Conjunto de datos activo - Seleccionar conjunto de datos activo...
 - | - Actualizar conjunto de datos activo
 - | - Ayuda sobre el conjunto de datos activo (si es posible)
 - | - Variables del conjunto de datos activo
 - | - Establecer nombres de casos...
 - | - Filtrar el conjunto de datos activo...
 - | - Apilar variables del conjunto de datos activo...
 - | - Eliminar los casos con valores omitidos...
 - | - Guardar el conjunto de datos activos...
 - | - Exportar el conjunto de datos activo...
- | - Modificar variables del conjunto de datos activo - Recodificar variables...
 - | - Calcular una nueva variable
 - | - Añadir números de observaciones al conjunto de datos
 - | - Tipificar variables...
 - | - Convertir variable numérica en factor...
 - | - Segmentar variable numérica...
 - | - Reordenar niveles de factor...
 - | - Definir contrastes de un factor...
 - | - Renombrar variables...
 - | - Eliminar variables del conjunto de datos...

Estadísticos - Resúmenes - Conjunto de datos activo

- | - Resúmenes Numéricos...
- | - Distribución de frecuencias...
- | - Número de observaciones ausentes
- | - Tabla de estadísticas...
- | - Matriz de correlaciones...
- | - Test de correlación
- | - Test de normalidad de Shapiro-Wilk...

- | - Tablas de contingencia - Tabla de doble entrada...
- |
 - | - Tabla de entradas múltiples...
 - | - Introducir y analizar una tabla de doble entrada...
- Medias - Test t para una muestra...
 - | - Test t para muestras independientes...
 - | - Test t para datos relacionados...
 - | - ANOVA de un factor...
 - | - ANOVA de múltiples factores...
- Proporciones - Test de proporciones para una muestra...
 - | - Test de proporciones para dos muestras...
- Varianzas - Test F para dos varianzas...
 - | - Test de Bartlett
 - | - Test de Levène
- Test no paramétricos - Test de Wilcoxon para dos muestras...
 - | - Test de Wilcoxon para muestras emparejadas...
 - | - Test de Kruskal-Wallis...
- Análisis dimensional - Fiabilidad de escala...
 - | - Análisis de componentes principales...
 - | - Análisis factorial...
 - | - Análisis de agrupación - Agrupación por k-medias...
 - | - Agrupación jerárquica...
 - | - Resumir la agrupación jerárquica...
 - | - Agregar la agrupación jerárquica al conjunto de datos...
- Ajuste de modelos - Regresión lineal...
 - | - Modelo lineal...
 - | - Modelo lineal generalizado...
 - | - Modelo logit multinominal...
 - | - Modelo ordinal de regresión...

Gráficas - Gama de colores

- | - Gráfica secuencial...
- Histograma...
- Gráfica de tallos y hojas...
- Diagrama de caja...
- Gráfica de comparación de cuantiles...
- Diagrama de dispersión...
- Matriz de diagramas de dispersión...
- Gráfica de líneas...
- Gráfica XY...
- Gráfica de las medias...
- Gráfica de barras...
- Gráfica de sectores...
- Gráficos 3D - Diagrama de dispersión en 3D...
 - | - Identificar observaciones con el ratón
 - | - Guardar gráfico en fichero...
- Guardar gráfico en fichero - como mapa de bits...
 - | - como PDF/Postscript/EPS...
 - | - Gráfica 3D RGL...

Modelos - Selecciona el modelo activo...

- Resumir el modelo
- Añadir las estadísticas de las observaciones a los datos...
- Intervalos de confianza...
- Test de hipótesis - Tabla ANOVA
 - Comparar dos modelos...
 - Hipótesis lineal...
- Diagnósticos Numéricos - Factores de inflación de varianza
 - Test de Breusch-Pagan para heteroscedasticidad...
 - Test de Durbin-Watson para autocorrelación...
 - Test RESET de no linealidad...
 - Test de valores atípicos de Bonferroni
- Gráficas - Gráficas básicas de diagnóstico
 - Gráfica de comparación de cuantiles de los residuos...
 - Gráfica de componentes + residuos
 - Gráfica de variables agregadas
 - Gráfica de influencia
 - Gráfica de los efectos

Distribuciones - Distribuciones continuas - distribución normal - Cuantiles normales...

- | | | | |
|--|--|---------------------------------------------------------|-----------------------------------------------|
| | | | - Probabilidades normales... |
| | | | - Gráfica de la distribución normal... |
| | | | - Muestra de una distribución normal... |
| | | - Distribución t - Cuantiles t... | |
| | | | - Probabilidades t... |
| | | | - Gráfica de la distribución t... |
| | | | - Muestra de una distribución t... |
| | | - Distribución Chi-cuadrado - Cuantiles Chi-cuadrado... | |
| | | | - Probabilidades Chi-cuadrado... |
| | | | - Gráfica de la distribución Chi-cuadrado... |
| | | | - Muestra de una distribución Chi-cuadrado... |
| | | - Distribución F - Cuantiles F... | |
| | | | - Probabilidades F... |
| | | | - Gráfica de la distribución F... |
| | | | - Muestra de una distribución F... |
| | | - Distribución exponencial - Cuantiles exponenciales... | |
| | | | - Probabilidades exponenciales... |
| | | | - Gráfica de la distribución exponencial... |
| | | | - Muestra de una distribución exponencial... |
| | | - Distribución uniforme - Cuantiles uniformes... | |
| | | | - Probabilidades uniformes... |
| | | | - Gráfica de la distribución uniforme... |
| | | | - Muestra de una distribución uniforme... |
| | | - Distribución beta - Cuantiles beta... | |
| | | | - Probabilidades beta... |
| | | | - Gráfica de la distribución beta... |
| | | | - Muestra de una distribución beta... |

	- Distribución de Cauchy - Cuantiles de Cauchy...
	- Probabilidades de Cauchy...
	- Gráfica de la distribución de Cauchy...
	- Muestra de una distribución de Cauchy...
	- Distribución logística - Cuantiles logísticos...
	- Probabilidades logísticas...
	- Gráfica de la distribución logística...
	- Muestra de una distribución logística...
	- Distribución lognormal - Cuantiles lognormal...
	- Probabilidades lognormal...
	- Gráfica de la distribución lognormal...
	- Muestra de una distribución lognormal...
	- Distribución gamma - Cuantiles gamma...
	- Probabilidades gamma...
	- Gráfica de la distribución gamma...
	- Muestra de una distribución gamma...
	- Distribución Weibull - Cuantiles Weibull...
	- Probabilidades Weibull...
	- Gráfica de la distribución Weibull...
	- Muestra de una distribución Weibull...
	- Distribución Gumbel - Cuantiles Gumbel
	- Probabilidades Gumbel...
	- Gráfica de la distribución Gumbel...
	- Muestra de una distribución Gumbel...
- Distribuciones discretas	- Distribución binomial - Cuantiles binomiales...
	- Probabilidades binomiales acumuladas...
	- Probabilidades binomiales...
	- Gráfica de la distribución binomial...
	- Muestra de una distribución binomial...
	- Distribución de Poisson - Cuantiles de Poisson...
	- Probabilidades de Poisson acumuladas...
	- Probabilidades de Poisson...
	- Gráfica de la distribución de Poisson...
	- Muestra de una distribución de Poisson...
	- Distribución geométrica - Cuantiles geométricos...
	- Probabilidades geométricas acumuladas...
	- Probabilidades geométricas...
	- Gráfica de la distribución geométrica...
	- Muestra de una distribución geométrica...
	- Distribución hipergeométrica - Cuantiles hipergeométricos...
	- Probabilidades hipergeométricas acumuladas...
	- Probabilidades hipergeométricas...
	- Gráfica de la distribución hipergeométrica...
	- Muestra de una distribución hipergeométrica...
	- Distribución binomial negativa - Cuantiles binomiales negativos...
	- Probabilidades binomiales negativas acumuladas...
	- Probabilidades binomiales negativas...
	- Gráfica de la distribución binomial negativa...
	- Muestra de una distribución binomial negativa...

Herramientas - Cargar paquete(s)...

- | - Cargar plug-in(s) de Rcmdr
- | - Opciones...

Ayuda - Ayuda de R Commander

- | - Introducción a R Commander
- | - Ayuda sobre el conjunto de datos activo (si es posible)
- | - Información sobre Rcmdr