

Manual sintético
para el uso del facilitador de R
versión 1

Vicente Manzano-Arrondo

- 2: [De qué va esto](#)
- 3: [Generar un archivo de datos](#)
- 6: [Antes de seguir instala R](#)
- 7: [El facilitador. Relación con R](#)
- 8: [Los archivos index.htm y DAD.R](#)
- 9: [La estructura del facilitador](#)
- 11: [Ejecutar el código en R](#)
- 12: [Tabla de frecuencias \(unitario\)](#)
- 14: [Medidas de representación y bondad \(unitario\)](#)
- 15: [Representaciones gráficas \(unitario\)](#)
- 18: [Inferencia unitaria: intervalo de confianza y PSHN](#)

De qué va esto

R es un potente recurso para realizar todo tipo de análisis de datos. Además de ello, se puede utilizar de forma gratuita y está siendo continuamente desarrollado por una comunidad internacional de personas cuyo objetivo es aumentar las posibilidades de análisis. No obstante, es un útil pensado para especialistas, que no buscan la belleza o la facilidad o el manejo de botones en opciones de menú, sino un alto nivel de control sobre lo que se hace y poder hacerlo todo (o casi). En otras palabras, R no es la mejor opción para las personas que no se ilusionan con los análisis de datos, sino que tienen otras motivaciones principales en la vida. En este grupo numeroso se encuentra el estudiantado universitario de ciencias sociales, humanas y de la salud. Necesitan hacer análisis de datos, pero desearían dedicarse a otras cosas.

Al otro lado estoy yo, un docente universitario que investiga por pasión y profesión y que enseña cómo investigar en psicología con las mismas motivaciones. Pero no soy un programador. Escribí algunos libros sobre SPSS y he ido probando diversos útiles informáticos, no solo para realizar los análisis de mis propias investigaciones, sino pensando especialmente en quienes vienen al aula para aprender de forma efectiva cómo hacer lo mismo. Tras unos años dudando, finalmente opté por decantarme por R. Mis estudiantes no se ven en la obligación de gastar dinero en ningún software, pero sufrirían lo espartano del sistema. Por eso he desarrollado una pequeña ayuda que, en plan doméstico, llamo *facilitador de R*.

Parto de la siguiente lógica: mis estudiantes...

1. ... saben utilizar una hoja de cálculo, como Calc de LibreOffice, por ejemplo. Si no saben manejar ninguna, es fácil y rápido solucionarlo y hacerlo por su cuenta. Este conocimiento, además, les resultará útil más allá de unas clases de análisis de datos. Gracias al manejo de una hoja de cálculo, se pueden escribir y guardar datos. Además, se pueden crear variables derivadas, manejando algunas expresiones y funciones sencillas.
2. ... navegan por Internet con el navegador que se les antoja, desde el sistema operativo que les place. No tienen problemas en este uso ni en las lógicas de manejo de formularios web.
3. ... van a enfrentarse a retos de análisis con características mucho más reducidas y menos variadas que las que justifican la gran diversidad de posibilidades de R. Luego, con algunas restricciones,

buena parte del código R puede ser muy repetitivo y, por tanto, facilitado de manera automática.

Si tales supuestos son ciertos, manejar R puede llegar a ser muy sencillo. Este manual sintético orienta sobre cómo hacerlo, siguiendo el programa concreto de una asignatura introductoria al análisis de datos en ciencias sociales, la que imparto en primer curso del grado de psicología. Es un contenido con muchas similitudes respecto a otros programas.

A grandes rasgos, el proceso sigue este esquema:

1. Mediante una hoja de cálculo, se genera el archivo de datos. El epígrafe “Generando el archivo de datos” orienta sobre ello.
2. Mediante el facilitador, se genera un archivo con el código R que se necesita para realizar los análisis deseados. El grueso de este manual, comenzando por “Cómo preparar R para su uso”, se dedica a este asunto.
3. Desde R, se carga el archivo de código y se ejecuta.
4. Desde un procesador de textos, se abre el archivo de resultados y se hacen las modificaciones de formato que se requieran para generar el informe final.

Generar un archivo de datos

Los datos pueden escribirse con cualquier procesador de textos. No obstante, resulta especialmente cómodo llevar esta tarea a cabo mediante una hoja de cálculo. A su vez hay muchos programas de ordenador que se encargan de esta tarea. Voy a partir de uno en concreto: Calc de LibreOffice. Si tú utilizas otro, como puede ser Excel de Microsoft, la gestión es lógicamente la misma, aunque puede variar algún pequeño detalle. Este no es un manual de hoja de cálculo, así que hay que tomar decisiones operativas como esta.

Vamos a partir de una situación real, donde se preguntó a un grupo de estudiantes por algunas variables diversas:

- Sexo: si es Mujer u Hombre.
- Vive: si vive en la Costa o en el interior, en cuyo segundo caso le preguntamos si se trata de un entorno Rural o Urbano.

- Placer: el grado de placer o satisfacción que le genera el uso del smartphone, en una escala de 1 (prácticamente ningún placer) a 7 (un muy alto nivel de placer o satisfacción).
- Depende: grado (también de 1 a 7) de dependencia que mantiene respecto al uso de su smartphone.
- Utilidades: número de utilidades web y de redes sociales a las que acude normalmente (whatsapp, dropbox, google+, facebook, instagram, etc.)

Sexo	Vive	Placer	Depende	Utilidades
Mujer	Costa	3	4	7
Mujer	Costa	7	7	9
Hombre	Costa	6	6	5
Mujer	Urbano	6	5	7
Mujer	Costa	7	6	9
Mujer	Costa	5	6	6
Mujer	Urbano	5	5	6
Mujer	Rural	6	7	5
Mujer	Rural	3	5	5
Mujer	Costa	4	6	11
Mujer	Rural	5	6	9
Hombre	Urbano	4	4	6
Hombre	Urbano	5	6	5
Mujer	Urbano	6	6	8
Mujer	Costa	4	4	6
Mujer	Costa	7	5	7
Mujer	Rural	4	5	7
Hombre	Urbano	5	7	9
Mujer	Rural	6	7	12
Mujer	Rural	2	5	6
Mujer	Urbano	7	7	10
Hombre	Urbano	6	7	9
Mujer	Costa	5	7	8
Mujer	Urbano	4	5	5
Mujer	Urbano	4	6	7

Tabla 1. Datos en Calc de LibreOffice

La tabla 1 contiene los resultados respecto a 25 estudiantes que facilitaron esta información. Un programa de hoja de cálculo permite dar formato, como puede ser el uso de colores y tamaños de letra. Aunque lo que nos interesa es su cualidad como datos, no hay por qué sufrir con el proceso. Un poco de color en la vida alegra la existencia. No obstante, el colorido y

cualquier otro aspecto del formato, se perderá cuando entremos en la espartana dimensión de los análisis de datos.

En este caso, todas las personas facilitaron información sobre todas las variables. Por ello, este conjunto de 5 variables x 25 casos no contiene casillas en blanco. En muchas ocasiones no ocurrirá así. No hay problema, se deja el hueco sin rellenar y punto. Para este ejemplo, además, no necesitamos crear nuevas variables. Si fuera el caso, utiliza tus conocimientos o consulta los manuales, tutoriales y demás recursos disponibles para aprender cómo hacerlo con tu hoja de cálculo.

Sexo,Vive,Placer,Depende,Utilidades
Mujer, Costa, 3, 4, 7
Mujer, Costa, 7, 7, 9
Hombre, Costa, 6, 6, 5
Mujer, Urbano, 6, 5, 7
Mujer, Costa, 7, 6, 9
Mujer, Costa, 5, 6, 6
Mujer, Urbano, 5, 5, 6
Mujer, Rural, 6, 7, 5
Mujer, Rural, 3, 5, 5
Mujer, Costa, 4, 6, 11
Mujer, Rural, 5, 6, 9
Hombre, Urbano, 4, 4, 6
Hombre, Urbano, 5, 6, 5
Mujer, Urbano, 6, 6, 8
Mujer, Costa, 4, 4, 6
Mujer, Costa, 7, 5, 7
Mujer, Rural, 4, 5, 7
Hombre, Urbano, 5, 7, 9
Mujer, Rural, 6, 7, 12
Mujer, Rural, 2, 5, 6
Mujer, Urbano, 7, 7, 10
Hombre, Urbano, 6, 7, 9
Mujer, Costa, 5, 7, 8
Mujer, Urbano, 4, 5, 5
Mujer, Urbano, 4, 6, 7

Cuadro 1. Archivo de datos.

Una vez tengas la carpeta de trabajo y hayas guardado en ella tu archivo de datos, se acabó el proceso con la hoja de cálculo. El resto es cosa del facilitador.

Antes de seguir, instala R

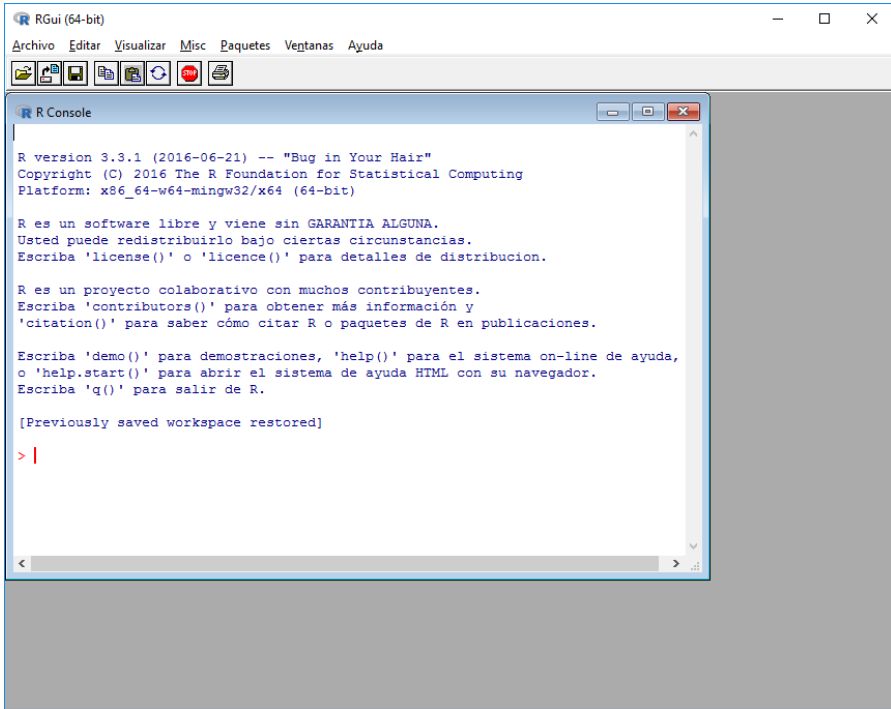


Figura 1. R se ejecutó con éxito.

Si navegas un poco por la red, verás diversas posibilidades para bajar e instalar R. Aquí te facilito la vía más ortodoxa, que consiste en acceder a la página oficial:

<http://cran.r-project.org/>

Busca descargar la última versión que se corresponda con tu sistema operativo. Tienes R para Windows, Linux o Mac OS. Por ejemplo, hay un enlace con el título “Download R for Windows”.

Es habitual que se aconseje descargar algo más que R. Por ejemplo, RStudio o R Commander. Para lo que vamos a hacer aquí, nada de ello es necesario. Si lo deseas, baja lo que quieras, investiga, juega, haz cálculos aritméticos, crea matrices, realiza análisis... por tu cuenta y cargando los paquetes que te apetezca. Hay multitud de manuales y tutoriales gratis en

la web, en formato web o en pdf. Yo no voy a suponer que has hecho nada de ello. Solo bajaste R y ya está.

Si el proceso ha sido exitoso, ejecuta el programa. Deberás obtener algo parecido a lo que ves en la figura 1. Si es así, enhorabuena. Ya podemos pasar a conocer de qué va el facilitador.

El facilitador. Relación con R

Algo que facilita no debe ser difícil. Obvio. Para que cumpla su cometido, el facilitador reduce drásticamente las posibilidades de R, situándose en lo más habitual, especialmente en relación al aprendizaje del análisis de datos. El primer ejemplo de lo que estoy diciendo se refiere a la estructura del archivo de datos que puede leer R. De todas esas posibilidades, el facilitador supone que (1) el archivo se encuentra en código txt o csv, (2) los datos están separados por comas, (3) cada línea es un caso, (4) la primera línea contiene los nombres de las variables, y (5) el punto decimal es un punto y no una coma. R permite más posibilidades. Pero si las quieres, deja de leer esto y pásate a un manual de R.

La lógica de uso, como ya he indicado, es que generas los datos con una hoja de cálculo, los guardas en formato de texto en una carpeta específica, entras en el facilitador, le indicas qué es lo que quieres hacer, el facilitador genera el conjunto de instrucciones R, capturas ese conjunto y lo guardas en disco desde tu procesador de textos o algo similar al block de notas de Windows, ejecutas R, le dices que lea ese archivo, y finalmente abres el resultado con tu procesador. Este proceso es fácil para una sesión de trabajo. Se vuelve tedioso si estás probando, si no tienes claro qué quieres, etc. pues en ese caso tendrás que repetirlo varias veces seguidas.

Bien, el facilitador se encuentra en esta dirección:

<http://asignatura.us.es/dadpsico/r/>

La primera vez necesitarás una conexión web. No será necesario en las siguientes, si bajas la utilidad y la instalas en tu ordenador. Para bajarla, pulsa en el enlace "Facilitador en zip", que se encuentra bajo el título de la página. Se trata de un archivo comprimido. Descárgalo donde quieras en tu ordenador, descomprímelo y ya estará listo para ser usado en las siguientes ocasiones.

Los archivos index.htm y DAD.R

Vamos a llamar *carpeta de programa* a la que tú has creado para descomprimir en ella al facilitador. En esa carpeta encontrarás:

- El archivo principal del facilitador: index.htm. Al ejecutarlo, se abre tu navegador y tienes ante ti el programa del que va este documento. Puedes buscarlo y ejecutarlo cada vez desde esa carpeta o bien idear otra estrategia. Una muy recomendable es incluir ese enlace en tu navegador, en la barra de enlaces o en la carpeta de favoritos.
- El archivo principal de funciones R: DAD.R. Contiene funciones necesarias para hacer los análisis.
- El archivo de datos en código texto que se utilizan en este manual: datosEjemplo.csv
- La carpeta /rec, con archivos javascript y de hojas de estilos, que gestionan el facilitador.

Vamos a llamar *carpeta de trabajo* a la que tú creas para colocar en ella el archivo de datos que te interese en cada momento y donde se guardarán los resultados de los análisis. Es importante organizarse bien. Mejor no saturar la *carpeta de programa* con datos y salidas. Mantenla intacta y crea una *carpeta de trabajo* para cada conjunto de datos que quieras analizar.

| *Importante: una vez creada una carpeta de trabajo, copia en ella DAD.R* |

El facilitador te pedirá, en cada sesión, que le indiques cuál va a ser la carpeta de trabajo. En ella irá generando los archivos con las salidas de los análisis y también los archivos que contienen las representaciones gráficas. Para hacer todo ello, el facilitador irá leyendo las instrucciones desde el archivo DAD.R, que cargará desde esa misma carpeta de trabajo. En otras palabras: cada vez que vayas a realizar análisis de datos con un nuevo archivo de datos:

1. Crea una carpeta específica.
2. Copia en ella tu archivo de datos en formato texto.
3. Copia en ella el archivo DAD.R

La estructura del facilitador

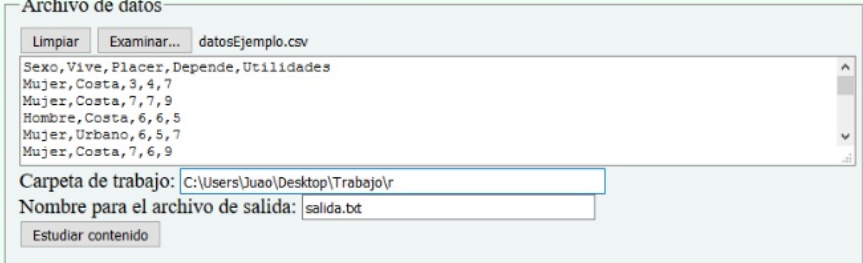
El facilitador está organizado en tres áreas. La primera tiene la apariencia que puedes ver en la figura 2.



The screenshot shows a window titled "Archivo de datos". At the top, there are two buttons: "Limpiar" and "Examinar...", followed by the text "datosEjemplo.csv". Below this is a large empty rectangular area. Underneath the empty area, there are two text input fields. The first is labeled "Carpeta de trabajo:" and is empty. The second is labeled "Nombre para el archivo de salida:" and contains the text "salida.txt". At the bottom left of the window is a button labeled "Estudiar contenido".

Figura 2. Área de datos

Pulsando el botón "Examinar", cargas en la ventana el contenido del archivo de datos. En el recuadro "Carpeta de trabajo", escribe la dirección completa de la carpeta de trabajo que hayas creado, y en donde se encuentra el archivo de datos. Cada sistema operativo tiene su criterio para mostrar esa información. En Windows, por ejemplo, basta con utilizar el explorador, hasta llegar a la carpeta de trabajo. Su nombre se muestra en la parte superior de la ventana. Sitúa el ratón sobre la carpeta, pulsa el botón derecho, y escoge la opción "Copiar dirección como texto". Después vuelve al facilitador y pulsa [Ctrl]+[V] o botón derecho / Pegar. Un tercer recuadro de texto es el del nombre que va a tener el archivo de salida. El facilitador sugiere "salida.txt". Modifícalo como quieras. Es bueno que siga manteniendo la extensión ".txt" para que sea más sencillo su reconocimiento por otros programas.



The screenshot shows the same "Archivo de datos" window, but now the large rectangular area contains a list of data. The list has a header row: "Sexo,Vive,Placer,Depende,Utilidades". Below the header are five rows of data: "Mujer,Costa,3,4,7", "Mujer,Costa,7,7,9", "Hombre,Costa,6,6,5", "Mujer,Urbano,6,5,7", and "Mujer,Costa,7,6,9". To the right of the list is a vertical scrollbar. Below the list, the "Carpeta de trabajo:" field now contains the path "C:\\Users\\Juao\\Desktop\\Trabajo\\r". The "Nombre para el archivo de salida:" field still contains "salida.txt". The "Estudiar contenido" button is still at the bottom left.

Figura 3. Área de datos con información completa.

La figura 3 muestra un aspecto completo de este área, una vez se han completado los pasos indicados. Según el navegador y la versión del sistema operativo, las barras inclinadas que separan las carpetas de las subcarpetas, pueden ser simples (/) o invertidas (\). No te preocupes, el facilitador se encarga de entender cualquiera de ambas.

La siguiente área, de opciones de análisis, es la más compleja, y contiene el meollo de las opciones de análisis. Su apariencia depende de lo que estés buscando en cada momento: análisis unitario, análisis relacional o aspectos de muestreo. La figura 4 muestra el estado para análisis unitarios. Su contenido concreto va a ser desvelado en secciones específicas, por lo que ahora basta con echarle un vistazo.

Opciones de análisis - UNITARIO -

El estudio unitario afecta a las variables:

Depende,Placer,Utilidades

Borrar la selección

Vive

Placer

Depende

Utilidades

Tabla de frecuencias con

☒ Información básica

☐ Acumulación

☐ Cálculos

☐ Z

☒ Intervalos de confianza

Medidas numéricas

☐ Moda y porcentaje

☐ Mediana y más

☐ Media y más

☐ Representación gráfica

Título:

Tipo de representación: Ciclograma

Inferencia

Trabajar con una seguridad del 95 %

Estadístico: media

☐ Construir un intervalo de confianza alrededor de la media

☐ Inferir si puede afirmarse que la media es .2 en la población, mediante una prueba de dos colas

Insertar código

Figura 4. Área de análisis

Por último, se encuentra el área de resultados o de código R. Al pulsar el botón "Insertar código" del área de análisis, se genera el listado de instrucciones que hay que ejecutar en R y se insertan en el área de resultados. La figura 5 muestra un ejemplo concreto.


```
Resultado
# 25 casos y 5 variables: Sexo, Vive, Placer, Depende, Utilidades.
# Carpeta y datos de trabajo:
setwd("C:/Users/Vicente/Desktop/Trabajo/DAD2016_2017/web/r")
datos = read.table("datosEjemplo.csv", header=TRUE, sep=";", fill=TRUE, strip.white=TRUE,
blank.lines.skip=TRUE)
attach(datos)
source("rec/DAD.R")
# Archivo de salida:
sink("sal/salida.txt",append=F)
#Tablas de frecuencias
cat("\nTabla de frecuencias de Placer:", "\n")
```

Limpiar

Figura 5. Área de código

Ejecutar el código en R

Una vez generado el código en el facilitador, se copia y se pega en R. Hay dos posibilidades para esto segundo. Lo más inmediato es copiar el código en la consola de R. La consola es la ventana que se abre al ejecutar R y que se observa en la figura 1. Al pegar el contenido del área de código del facilitador, se van ejecutando línea a línea como si se hubieran tecleado desde la propia consola. Esta es la mejor opción si no deseas saber nada de R, salvo que está ejecutando las órdenes, y punto.

Otra posibilidad es abrir una ventana de script y copiar ahí el código. Para abrir una ventana de script, sigue las opciones de menú Archivo / Nuevo script. El resultado se muestra en la figura 6. Copia en ella el contenido del área de código del facilitador. A partir de aquí, caben varias posibilidades. Una de ellas, similar a copiar el código en la ventana de consola, es ordenar a R que lo ejecute todo. Se consigue con Editar / Ejecutar todo. Es más interesante poner en práctica otra cosa. Observa los botones de la barra superior de R. Uno de los botones tiene esta apariencia: 

Al pulsar ese botón, se ejecuta la línea de código del script, en donde se encuentre en ese momento el cursor. Esta forma de ejecutar las instrucciones es más lenta, pero te permite observar qué ocurre en cada caso, identificar dónde se encuentra un error, caso de ocurrir, y de paso aprender algo de R sobre la marcha. Otra posibilidad es pulsar [Ctrl]+[R] y se genera el mismo efecto. Esta opción tiene un valor añadido: ejecuta el código resaltado o seleccionado. Si no hay ninguno, ejecuta la línea de código en donde se encuentre el cursor. Luego, otra forma de controlar el proceso de ejecución es capturar unas líneas de código y pulsar esa combinación de teclas, para provocar que lleve a cabo la porción de código

que interese en cada caso. Imagina, por ejemplo, que has pedido una tabla de frecuencias con una variable categórica, y pediste en el facilitador que la tabla incluyera la columna de distancias estandarizadas. Para ello, R debe calcular la media aritmética, entre otros estadísticos. Al intentarlo, se generará un mensaje de error, puesto que no es posible calcular la media con un listado de nombres en lugar de un conjunto de números. Al ejecutar todo el código completo y de golpe, esto se te puede escapar.

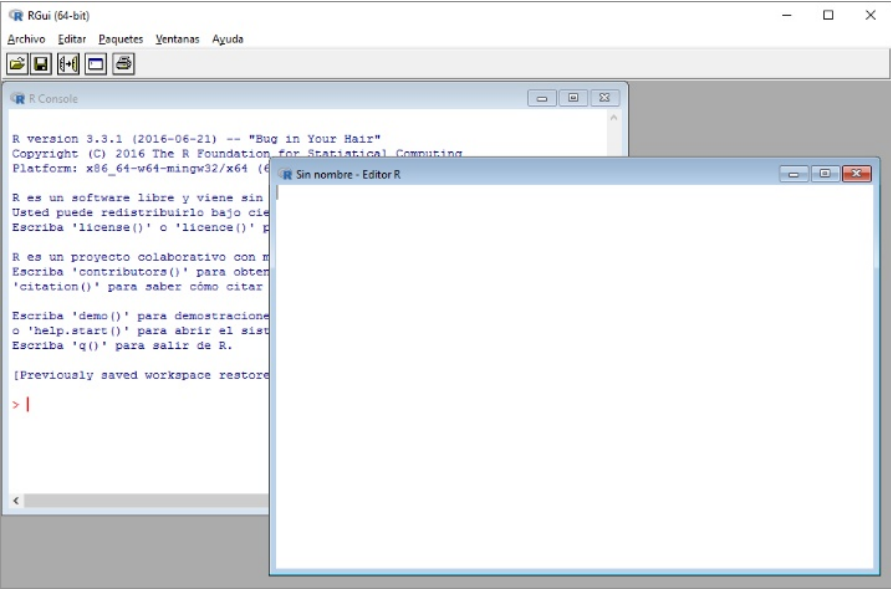


Figura 6. Vista de script en R.

Tabla de frecuencias (unitario)

```
# 25 casos y 5 variables: Sexo, Vive, Placer, Depende, Utilidades.
# Carpeta y datos de trabajo:
setwd("C:/Users/Joao/Desktop/Trabajo/DAD/web/r")
datos = read.table("datosEjemplo.csv", header=TRUE, sep=",",
fill=TRUE, strip.white=TRUE, blank.lines.skip=TRUE)
attach(datos)
source("DAD.R")
# Archivo de salida:
sink("salida.txt",append=F)
```

Cuadro 2. Cabecera de código

El primer ejercicio completo va a consistir en generar una tabla de frecuencias. Para ello, recorre todo el proceso previo: en el área de datos, carga el archivo `datosEjemplo.csv`, indica la carpeta de trabajo y decide el nombre para el archivo de salida. Al pulsar "Estudiar contenido", la ventana de código muestra el resultado del cuadro 2. Ese código le indica a R cuál es el archivo de datos, lo carga en memoria, carga también las funciones de DAD.R e indica al programa en qué archivo de texto ir escribiendo los resultados.

Pues bien, vamos ahora a la primera parte del área de análisis, donde se encuentra activo el modo UNITARIO. Escoge la variable "Placer" en el listado de variables del área de análisis unitario. El recuadro "El estudio unitario afecta a las variables:" muestra el texto "Placer". Si te equivocas seleccionando variables, pulsa "Borrar selección" y comienza.

El área específica para tabla de frecuencias cuenta con la opción "Información básica". Debe estar seleccionada para que se genera la tabla y provoca tres columnas: valores, frecuencias absolutas (f) y porcentajes (p). La casilla "Acumulación", añade otras dos columnas: frecuencias absolutas acumuladas (Fa) y porcentajes acumulados (Pa). Obviamente, esto solo tiene sentido si la variable es al menos ordinal. La casilla "Cálculos", inserta dos columnas: valores por frecuencias (xf) y valores al cuadrado por sus frecuencias (x2f), resultados necesarios para los cálculos a mano de la media y la varianza. La casilla "Z" inserta una columna con las distancias estandarizadas que se corresponden con cada uno de los valores. Y la casilla "Intervalos de confianza" inserta dos columnas, con los límites inferior (Linf) y superior (Lsup) de un intervalo de confianza del 95% en torno a las proporciones de cada valor.

```
#Tablas de frecuencias
cat("\nTabla de frecuencias de Placer:", "\n")
fTabFrec (Placer, acu=T, cal=T, zet=T, int=T)
```

Cuadro 3a. Código para una tabla de frecuencias.

Si no hay nada más en el área de análisis que esté seleccionado, y has seleccionado todas las casillas de la sección de tabla de frecuencias, entonces al pulsar el botón "Insertar código", la ventana de código debe mostrar la información del cuadro 3a a continuación del código que ya estaba escrito en esa ventana. Al ejecutarlo, el archivo de salida (sea "salida.txt" o el que hayas indicado al facilitador) contiene lo que puedes ver en el cuadro 3b.

Tabla de frecuencias de Placer:										
	f	p	Fa	Pa	xf	x2f	Z	Linf	Lsup	
2	1	4	1	4	2	4	-2.221	0.0	11.7	
3	2	8	3	12	6	18	-1.490	0.0	18.6	
4	6	24	9	36	24	96	-0.760	7.3	40.7	
5	6	24	15	60	30	150	-0.029	7.3	40.7	
6	6	24	21	84	36	216	0.701	7.3	40.7	
7	4	16	25	100	28	196	1.432	1.6	30.4	
Tot	25	100	NA	NA	126	680	NA	NA	NA	

Cuadro 3b. Tabla de frecuencias completa para la variable Placer.

Observa el resultado. La fila de totales contiene algunos "NA". Es un elemento R que implica la ausencia de valor. Esos espacios no significan nada, porque no tiene sentido sumar esas columnas.

Medidas de representación y bondad (unitario)

La segunda sección del modo UNITARIO muestra las opciones para generar medidas de representación (moda, mediana y media), con sus respectivas medidas de bondad: porcentaje para la moda; mediana de distancias a la mediana, para la mediana; y desviación tipo y desviación acotada, para la media. A su vez, si has solicitado mediana o media, al suponer que los datos pueden por lo menos ordenarse, el facilitador incluye también los valores mínimo y máximo. En todos los casos, la tabla resultante muestra el número de casos considerado.

```
# Descriptivos
cat("\nTabla de descriptivos:", "\n")
Variable = list("Placer", "Depende", "Utilidades")
.noUsar=rbind(fDescribe(Placer, mod=T, mdn=T, med=T))
.noUsar=rbind(.noUsar, fDescribe(Depende, mod=T, mdn=T, med=T))
.noUsar=rbind(.noUsar, fDescribe(Utilidades, mod=T, mdn=T, med=T))
cbind(Variable, .noUsar)
```

Cuadro 4a. Código de medidas numéricas para modo unitario.

Tabla de descriptivos:												
	Variable	Moda	PModa	Mediana	MDM	Media	Dtip	Daco	Mín	Máx	n	
[1,]	"Placer"	4	24	5	1	5.04	1.34	55	2	7	25	
[2,]	"Depende"	6	32	6	1	5.76	0.99	67	4	7	25	
[3,]	"Utilidades"	5	21	7	2	7.42	1.96	59	5	12	24	

Cuadro 4b. Tabla de medidas numéricas en modo unitario.

Si has seleccionado más de una variable, el resultado las agrupará en una tabla única. Los cuadros 4a y 4b muestran el código y el resultado que se generan al seleccionar las variables Placer, Depende y Utilidades, solicitar todas las medidas numéricas y ejecutar el código en R.

Por ejemplo, la variable "Utilidades" cuenta con una moda de valor 5, que se corresponde con el 21% de los datos ($n = 24$); la mediana es 7, con una mediana de distancias a ella de valor 2; la media es 7,42, la desviación típica es 1,96 y la acotada es del 59%; el mínimo y el máximo son, respectivamente, 5 y 12.

Si la variable es categórica, solo procede la moda. Es más, R mostrará todos los datos como si fueran literales. Los cuadros 5a y 5b muestran un ejemplo para la variable "Vive", observando que el valor más frecuente (40%) es la zona "Rural".

```
# Descriptivos
cat("\nTabla de descriptivos:", "\n")
Variable = list("Vive")
.noUsar=rbind(fDescribe(Vive,mod=T))
cbind(Variable,.noUsar)
```

Cuadro 5a. Código para medidas numéricas categóricas.

Tabla de descriptivos:				
	Variable	Moda	PModa	n
[1,]	"Vive"	"Costa"	"40"	"25"

Cuadro 5b. Resultado para medidas numéricas categóricas.

Representaciones gráficas (unitario)

Uno de los aspectos que más conquistan de R son sus amplias posibilidades para la generación de representaciones gráficas. El facilitador no entra en ello. Es más, para esta versión 1, el conjunto de posibilidades es reducido, pero operativo. Se consideran cinco tipos: ciclograma o diagrama de pastel o círculo, diagrama de barras, histograma, diagrama de caja y patillas para una variable, y diagrama de caja y patillas combinado para un conjunto de variables.

La gestión de esta variabilidad es simple en el facilitador. Consiste en (1) seleccionar la variable o variables sobre las que deseas que se genera la gráfica [cuidado: recuerda comprobar en el recuadro de variables seleccionadas, que solo está la que te interesa], (2) poner un título al

gráfico, (3) seleccionar el tipo de gráfico, y (4) pulsar "Insertar código". El cuadro 6 muestra el código que se genera para cada una de las siguientes cinco peticiones:

1. Ciclograma para Sexo (figura 7)
2. Diagrama de barras para Vive (figura 8)
3. Histograma para Utilidades (figura 9)
4. Diagrama de caja y patillas pra Utilidades (figura 10)
5. Diagrama de caja y patillas conjunto para las variables Placer, Depende y Utilidades (figura 11)

```
#Representaciones gráficas
sink()
jpeg("imagen1.jpg")
pie(table(Sexo),main="Sexo")
dev.off()
cat("\nSe ha creado el archivo de imagen \"imagen1.jpg\" con la
variable Sexo\\\"\\n\", \"\\n\")
sink("sal.txt", append=T)

#Representaciones gráficas
sink()
jpeg("imagen2.jpg")
barplot(table(Vive),main="Zona de residencia")
dev.off()
cat("\nSe ha creado el archivo de imagen \"imagen2.jpg\" con la
variable Vive\\\"\\n\", \"\\n\")
sink("sal.txt", append=T)

#Representaciones gráficas
sink()
jpeg("imagen3.jpg")
hist(Utilidades,main="Número de utilidades virtuales cotidianas")
dev.off()
cat("\nSe ha creado el archivo de imagen \"imagen3.jpg\" con la
variable Utilidades\\\"\\n\", \"\\n\")
sink("sal.txt", append=T)

#Representaciones gráficas
sink()
jpeg("imagen4.jpg")
boxplot(Utilidades,main="Número de utilidades virtuales cotidianas")
dev.off()
cat("\nSe ha creado el archivo de imagen \"imagen4.jpg\" con la
variable Utilidades\\\"\\n\", \"\\n\")
sink("sal.txt", append=T)
```



```
#Representaciones gráficas
sink()
jpeg("imagen5.jpg")
x=list(Placer,Depende,Utilidades)
boxplot(x,main="Placer, Depende y Utilidades")
dev.off()
cat("\nSe ha creado el archivo de imagen \"imagen5.jpg\"\n","\n")
sink("sal.txt",append=T)
```

Cuadro 6. Código de generación de gráficas unitarias.

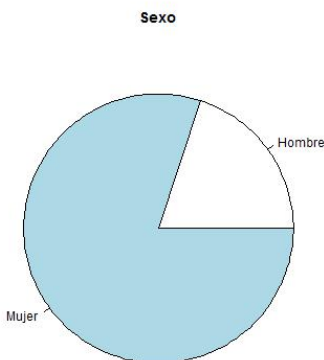


Figura 7. Ciclograma



Figura 8. Diagrama de barras

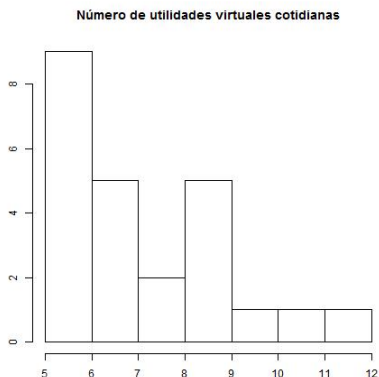


Figura 9. Histograma

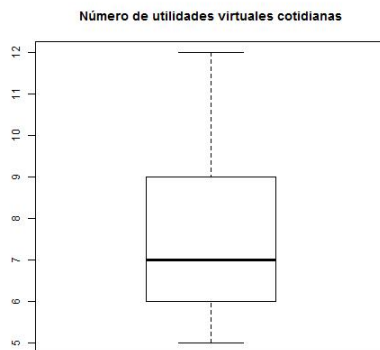


Figura 10: Caja y patillas

A cada cual le puse el título de gráfico que me pareció. El facilitador guarda las gráficas en formato jpeg y en archivos con el nombre imagenX.jpg, donde X es un número entero correlativo. Cada vez que

pulsas el botón "Estudiar contenido" del área de datos, el facilitador pone el contador de nuevo a 1. Si R encuentra un archivo con ese nombre, lo sustituye por el nuevo. Así, por ejemplo, el histograma para Utilidades tendrá el nombre imagen3.jpg. Las figuras se guardan en la *carpeta de trabajo*.

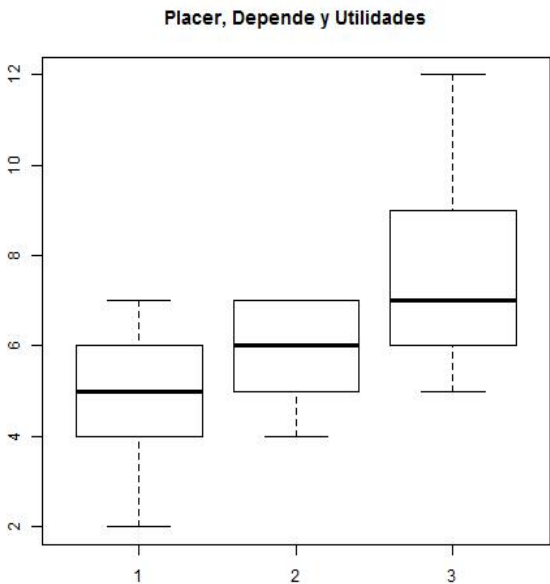


Figura 11. Caja y patillas conjunto para tres variables

Inferencia unitaria: intervalo de confianza y PSHN

Realizar una estimación por intervalo de una media o de una proporción, o poner en marcha una prueba de significación de la hipótesis nula, son tareas muy sencillas con el facilitador. Basta con ir pulsando las opciones que correspondan en cada uno de los casos mencionados. El propio facilitador se adapta a lo que vayas pulsando.

Por ejemplo, vamos a pedir un intervalo de confianza para la media aritmética de Placer, Depende y Utilidades, con un nivel de seguridad del 97%. Los cuadros 7a y 7b muestran el código y el resultado. Así, por ejemplo, la variable Depende cuenta con una media aritmética de valor 5,76 en la muestra y se puede afirmar que la media poblacional se encuentra en el intervalo {5,32 ; 6,20} con una confianza del 97%.

```
#Inferencia estadística
cat("\nTabla de inferencia:", "\n")
Variable = list("Placer", "Depende", "Utilidades")
.noUsar=rbind(inf.Medias(Placer,int=T,seg=0.97))
.noUsar=rbind(.noUsar,inf.Medias(Depende,int=T,seg=0.97))
.noUsar=rbind(.noUsar,inf.Medias(Utilidades,int=T,seg=0.97))
cbind(Variable,.noUsar)
```

Cuadro 7a. Código para intervalo de confianza

Tabla de inferencia:

	Variable	Media	Dtip	n	Conf	Linf	Lsup
[1,]	"Placer"	5.04	1.34	25	0.97	4.45	5.63
[2,]	"Depende"	5.76	0.99	25	0.97	5.32	6.2
[3,]	"Utilidades"	7.42	1.96	24	0.97	6.53	8.3

Cuadro 7b. Tabla de intervalos de confianza para tres variables

Para una prueba de significación de la hipótesis nula, vamos a ejemplificar el caso de una proporción. Probemos si puede afirmarse que la proporción de personas que viven en la costa es del 30%, con una seguridad del 95%, en una prueba de dos colas. Los cuadros 8a y 8b se encargan de ello.

```
#Inferencia estadística
cat("\nTabla de inferencia: [valor= Costa]", "\n")
Variable = list("Vive")
.noUsar=rbind(inf.Prop(Vive,val="Costa",hip=T,pi=.3))
cbind(Variable,.noUsar)
```

Cuadro 8a. Código para PSHN con una proporción

Tabla de inferencia: [valor= Costa]

	Variable	Frec	Prop	n	pi	Z	p2col
[1,]	"Vive"	10	0.4	25	0.3	1.021	0.3074

Cuadro 8b. Resultado de una PSHN para una proporción