

Econometría Aplicada con



```
R Console (32-bit)
Archivo Editar Misc. Ejecutar Ventanas Ayuda

> x <- c(1,2,3,4,5,6)
> y <- x^2
> print(y)
[1] 1 4 9 16 25 36
> mean(y)
[1] 15.16667
> var(y)
[1] 178.9444
> lm_1 <- lm(y ~ x)
> print(lm_1)

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept) -9.3333
x              7.0000

> summary(lm_1)

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept) -9.3333
x              7.0000

Residuals:
1      2      3      4      5      6
3.3333 -0.6667 -2.6667 -2.6667 -0.6667  3.3333

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -9.3333    2.8441   -3.282 0.030453 *
x              7.0000    0.7303    9.585 0.000662 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.055 on 4 degrees of freedom
Multiple R-squared:  0.9583,    Adjusted R-squared:  0.9478
F-statistic: 91.87 on 1 and 4 DF,    p-value: 0.000662

> |
```



SESIÓN 3: Exploración de datos

Contenido

Introducción	4
Ingresar datos desde el teclado	5
Importar datos	7
Importar datos de texto	7
Desde Rcmdr	8
RStudio	9
Importar datos de Excel	10
Desde Rcmdr	11
Desde RStudio	12
Importar datos de Stata	13
Rcmdr	13
RStudio	14
Exportar datos en R	16
Exportar a CSV	16
Exportar a texto separado por tabulaciones	16
Exportar datos a Excel	16
Exportar datos a Stata	16
Bibliografía	17
Recursos informáticos	17

Introducción

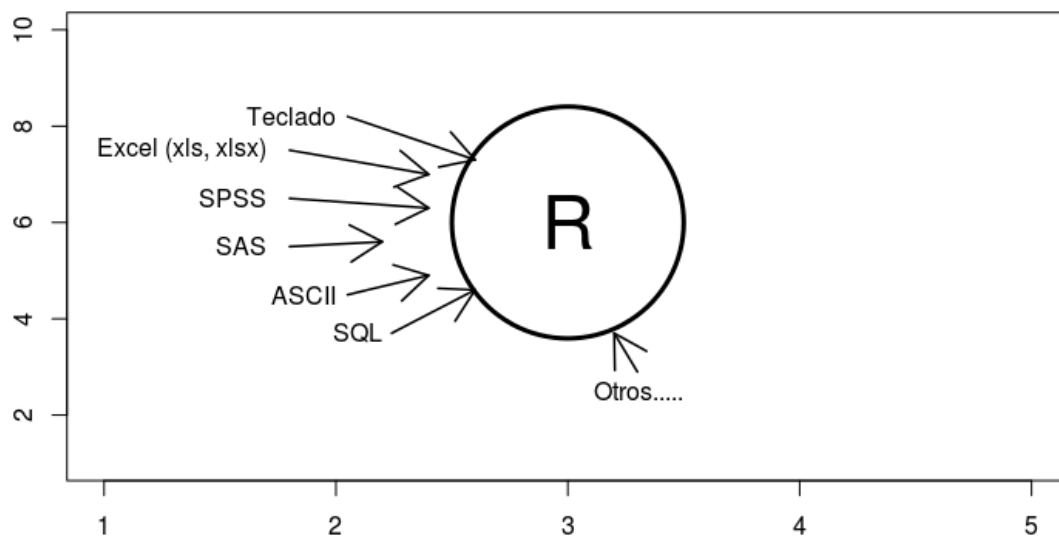
Para el análisis es necesario el trabajar con datos, para esto se los debe ingresar o importar si es que están en otros formatos, R acepta la importación de casi todos los formatos de almacenamiento de datos para el análisis, esto demuestra la versatilidad del mismo.

En la siguiente sesión se explicarán las diversas formas que el programa ofrece para la introducción de los datos, también se verán las formas que el mismo otorga para la exportación de datos en diferentes formatos.

Ingresar datos desde el teclado

Es sabido que R es compatible con casi todos los formatos de datos, por lo que se entiende que no existen problemas de compatibilidad a la hora de trabajar con datos de otras extensiones en R.

Entrada de datos



Para ingresar datos desde el teclado se debe crear un **dataframe**, este será el conjunto de datos, la sintaxis en ejemplo será la siguiente:

```
> df=data.frame(edad=numeric(0), sexo=character(0), peso=numeric(0))
```

Esto le indicará a R que se está creando un conjunto de datos llamado **df** con tres variables (edad, sexo y peso), que son de tipo numérico, en caracteres y numérico, respectivamente.

Para ingresar datos desde el teclado se debe usar el comando `edit()`:

```
> edit(df)
[1] edad sexo peso
```

Esto abrirá una ventana similar a una hoja de cálculo donde se podrán ingresar los datos de forma manual, directamente desde el teclado:

R Editor de datos

Archivo Editar Ayuda

	edad	sexo	peso	var4	var5	var6	var7
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							

En la ventana que aparece se podrán ingresar los datos de forma directa y se podrán guardar dichos datos:

R Editor de datos

Archivo Editar Ayuda

	edad	sexo	peso
1	12	H	50
2	13	M	62
3	12	M	63

Importar datos

Importar datos de texto

La función más importante para importar datos, o al menos una de las más usadas es **read.table()**, automáticamente convierte los datos en un dataframe.

Suponiendo que se tiene una base de datos de nombre **bandes.txt**, que es un archivo de texto con tabuladores que separan a las variables, se hará el siguiente procedimiento.

```
> Dataset = read.table("C:/Users/JNavarro/Downloads/band_des.txt",  
  header=TRUE, sep="\t", na.strings="NA", dec=".", strip.white=TRUE)
```

Este procedimiento indica lo siguiente:

- Primero la ubicación del archivo, que es un archivo de texto.
- Luego indica que en la primera fila están los nombres de las variables en el conjunto de datos.
- Los separadores para cada columna son tabuladores.
- Los valores perdidos se leerán como NA.
- El separador decimal de los números serán los puntos(.).

Para visualizar los datos se debe usar el comando **edit()**:

```
> edit(Dataset)
```

Aparecerá una ventana que se mostrará la base de datos:

R Editor de datos

Archivo Editar Ayuda

	DEPARTAMENTO	band2006	band2007	band2008	band2009	band2010	band2011
1	Amazonas	32	8	18	2	4	4
2	Áncash	10	27	42	22	105	64
3	Apurímac	NA	0	2	0	9	2
4	Arequipa	32	73	54	65	206	163
5	Ayacucho	37	23	31	43	41	17
6	Cajamarca	28	9	9	15	18	43
7	Callao	7	8	22	18	28	16
8	Cusco	2	21	0	237	283	532
9	Huancavelica	NA	0	19	3	25	2
10	Huánuco	38	27	30	7	14	42
11	Ica	31	28	8	39	55	32
12	Junín	21	17	13	10	33	44
13	La Libertad	115	44	83	171	122	17
14	Lambayeque	52	58	0	143	195	122
15	Lima	612	477	842	806	944	821
16	Loreto	5	1	4	11	39	115
17	Madre de Dios	NA	1	0	9	4	1
18	Moquegua	6	7	7	6	3	5
19	Pasco	5	4	3	10	4	0

Desde Rcmdr

Rcmdr permite importar mediante botones, para eso se debe ir a **Datos > Importar Datos > Desde texto, portapapeles o URL**. Aparecerá una ventana en donde se deberá indicar el nombre del conjunto de datos, el indicador de valores perdidos, los separadores de campos y los separadores decimales:

R Leer archivo de texto, portapapeles o URL

Introducir el nombre del conjunto de datos:

Nombres de las variables en el archivo: ☒

Indicador de datos ausentes:

Localización del archivo de datos

☒ Sistema de archivo local

☐ Portapapeles

☐ Dirección URL

Separador de campos

☒ Espacio en blanco

☐ Comas




☐ Tabuladores

☐ Otro Especificar:

Carácter decimal

☒ Punto [.]

☐ Coma [,]

 Ayuda  Aceptar  Cancelar

Se buscará el conjunto de datos y se abrirá el conjunto de datos haciendo clic en el botón **Visualizar conjunto de datos**.

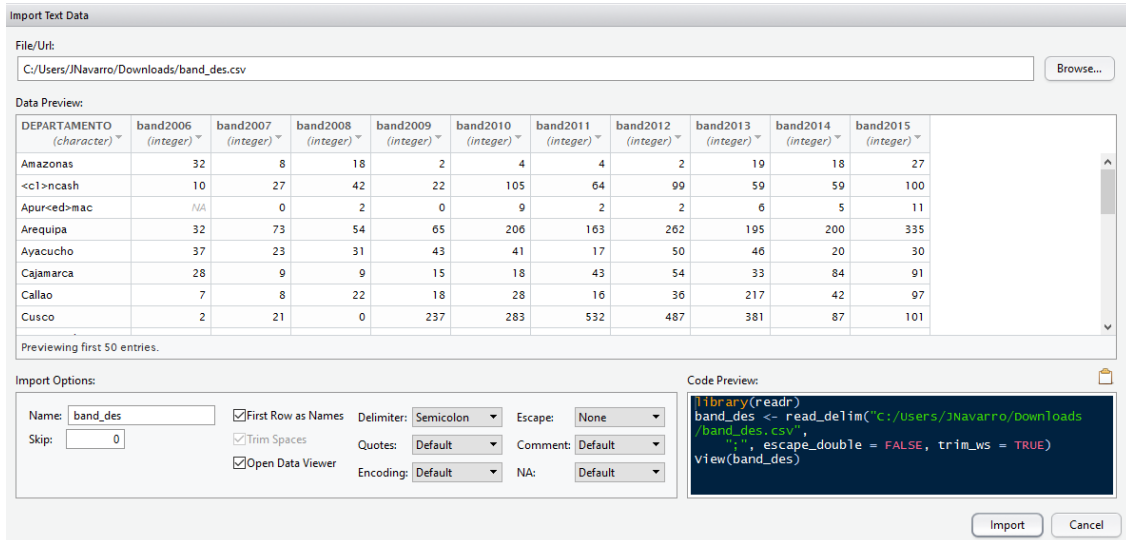
R Dataset

	DEPARTAMENTO	band2006	band2007	band2008	band2009	band2010
1	Amazonas	32	8	18	2	4
2	Áncash	10	27	42	22	105
3	Apurímac	NA	0	2	0	9
4	Arequipa	32	73	54	65	206
5	Ayacucho	37	23	31	43	41
6	Cajamarca	28	9	9	15	18
7	Callao	7	8	22	18	28
8	Cusco	2	21	0	237	283
9	Huancavelica	NA	0	19	3	25
10	Huánuco	38	27	30	7	14
11	Ica	31	28	8	39	55
12	Junín	21	17	13	10	33
13	La Libertad	115	44	83	171	122
14	Tumbes	52	52	0	142	105

RStudio

En RStudio existe una forma directa de importar archivos con extensión *.csv, para eso se debe ir en la parte superior derecha al botón **Import**

Dataset > From CSV, esto abrirá una ventana donde se deberá indicar y buscar el archivo que se va a importar:



Import Text Data

File/Url:
C:/Users/JNavarro/Downloads/band_des.csv

Data Preview:

DEPARTAMENTO (character)	band2006 (integer)	band2007 (integer)	band2008 (integer)	band2009 (integer)	band2010 (integer)	band2011 (integer)	band2012 (integer)	band2013 (integer)	band2014 (integer)	band2015 (integer)
Amazonas	32	8	18	2	4	4	2	19	18	27
<cl>ncash	10	27	42	22	105	64	99	59	59	100
Apur<ed>mac	NA	0	2	0	9	2	2	6	5	11
Arequipa	32	73	54	65	206	163	262	195	200	335
Ayacucho	37	23	31	43	41	17	50	46	20	30
Cajamarca	28	9	9	15	18	43	54	33	84	91
Callao	7	8	22	18	28	16	36	217	42	97
Cusco	2	21	0	237	283	532	487	381	87	101

Previewing first 50 entries.

Import Options:

Name: band_des
Skip: 0

☒ First Row as Names
☒ Trim Spaces
☒ Open Data Viewer

Delimiter: Semicolon
Escape: None
Quotes: Default
Comment: Default
Encoding: Default
NA: Default

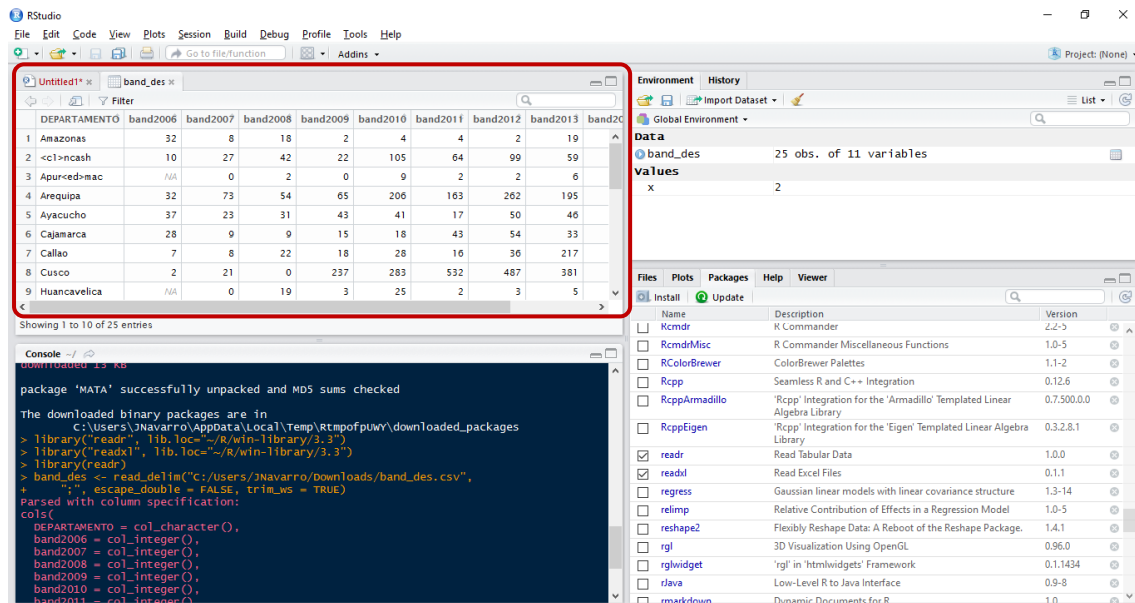
Code Preview:

```
library(readr)
band_des <- read_delim("C:/Users/JNavarro/Downloads/
band_des.csv",
  ";", escape_double = FALSE, trim_ws = TRUE)
view(band_des)
```

Import Cancel

Se debe indicar en esa tabla también las opciones que están en la parte inferior, como los delimitadores, la codificación, entre otros.

En la interfaz de RStudio se mostrarán los datos importados:



RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

band_des

DEPARTAMENTO	band2006	band2007	band2008	band2009	band2010	band2011	band2012	band2013	band2014	band2015
1 Amazonas	32	8	18	2	4	4	2	19	18	27
2 <cl>ncash	10	27	42	22	105	64	99	59	59	100
3 Apur<ed>mac	NA	0	2	0	9	2	2	6	5	11
4 Arequipa	32	73	54	65	206	163	262	195	200	335
5 Ayacucho	37	23	31	43	41	17	50	46	20	30
6 Cajamarca	28	9	9	15	18	43	54	33	84	91
7 Callao	7	8	22	18	28	16	36	217	42	97
8 Cusco	2	21	0	237	283	532	487	381	87	101
9 Huancavelica	NA	0	19	3	25	2	3	5		

Showing 1 to 10 of 25 entries

Console

```
package 'MATA' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
C:/Users/JNavarro/AppData/Local/Temp/Rtmpofpuwv/downloaded_packages
> library("readr", lib.loc = "~/R/win-library/3.3")
> library("readxl", lib.loc = "~/R/win-library/3.3")
> library(readr)
> band_des <- read_delim("C:/Users/JNavarro/Downloads/band_des.csv",
+ ";", escape_double = FALSE, trim_ws = TRUE)
Parsed with column specification:
col5(
  DEPARTAMENTO = col_character(),
  band2006 = col_integer(),
  band2007 = col_integer(),
  band2008 = col_integer(),
  band2009 = col_integer(),
  band2010 = col_integer(),
  band2011 = col_integer(),
  band2012 = col_integer(),
  band2013 = col_integer(),
  band2014 = col_integer(),
  band2015 = col_integer())
```

Environment

Global Environment

band_des 25 obs. of 11 variables

Files

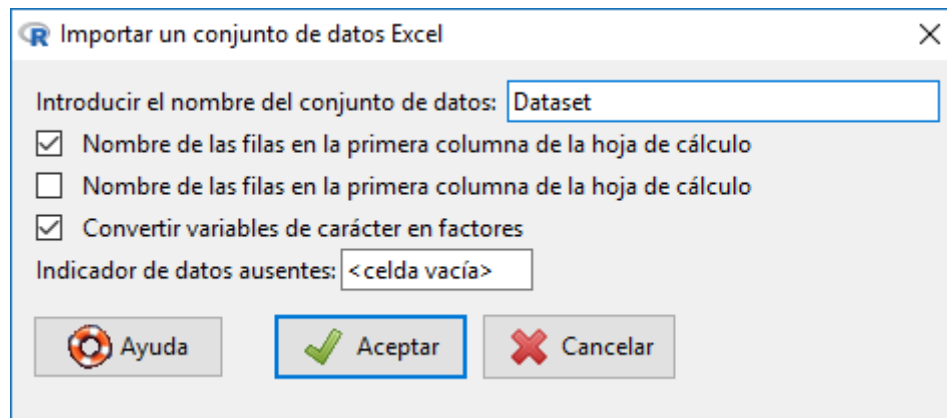
Name	Description	Version
KcMdr	R Commander	2.2-5
RcmdrMisc	R Commander Miscellaneous Functions	1.0-5
RColorBrewer	ColorBrewer Palettes	1.1-2
Rcpp	Seamless R and C++ Integration	0.12.6
RcppArmadillo	'Rcpp' Integration for the 'Armadillo' Templated Linear Algebra Library	0.7.500.0.0
RcppEigen	'Rcpp' Integration for the 'Eigen' Templated Linear Algebra Library	0.3.2.8.1
readr	Read Tabular Data	1.0.0
readxl	Read Excel Files	0.1.1
regress	Gaussian linear models with linear covariance structure	1.3-14
relimp	Relative Contribution of Effects in a Regression Model	1.0-5
reshape2	Flexibly Reshape Data: A Reboot of the Reshape Package.	1.4.1
rgl	3D Visualization Using OpenGL	0.96.0
rglwidget	'rgl' in 'htmlwidgets' Framework	0.1.1434
rJava	Low-Level R to Java Interface	0.9-8
rmarkdown	Dynamic Documents for R	1.0

Importar datos de Excel

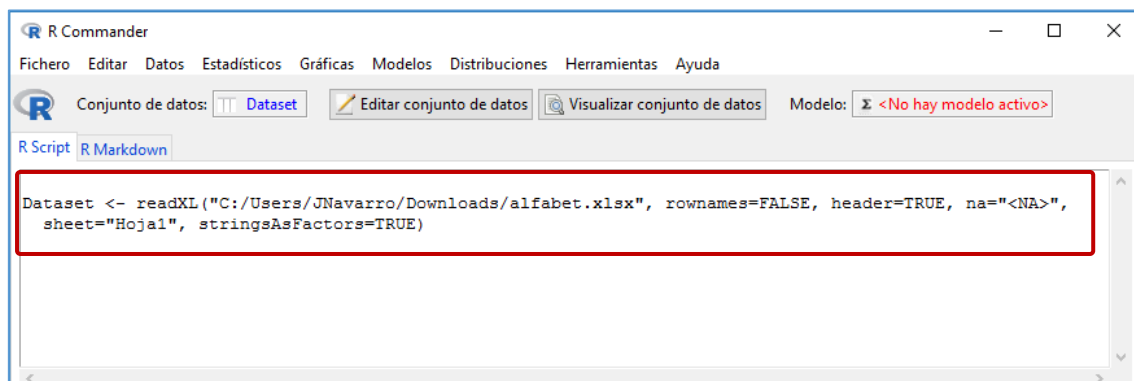
Lo mejor para leer ficheros Excel es guardar éstos en formato csv, es decir, texto delimitado por comas. También podemos leer el xls directamente con **read.excel()** del paquete **readxl**.

Desde Rcmdr

Se pueden importar datos desde este programa yendo a **Datos > Importar Datos > Desde un archivo de Excel**. Aparecerá una ventana en donde se indicará el nombre para el archivo de datos en R, si se quiere que las filas en la primera columna contengan a los nombres de las variables y también se especificará el nombre para los datos perdidos:



Luego de eso se abrirá el explorador donde se indicará la base de datos y el programa mostrará en el Script la sintaxis que importa a dichos datos.

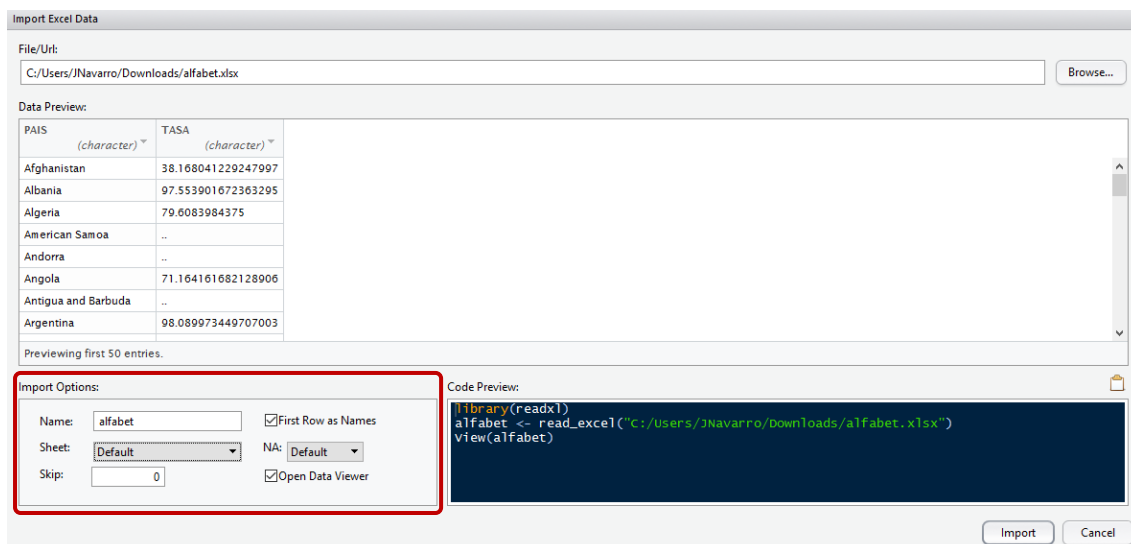


Para visualizar los datos importados se debe dar clic en el botón de **visualizar conjunto de datos**.

	PAIS	TASA
1	Afghanistan	38.168041229247997
2	Albania	97.553901672363295
3	Algeria	79.6083984375
4	American Samoa	..
5	Andorra	..
6	Angola	71.164161682128906
7	Antigua and Barbuda	..
8	Argentina	98.089973449707003
9	Armenia	99.768417358398395
10	Aruba	97.524200439453097
11	Australia	..
12	Austria	..
13	Azerbaijan	99.805259704589801
14	Bahamas, The	..
15	Bahrain	95.717262268066406

Desde RStudio

RStudio ofrece una forma rápida de importado de la base de datos de Excel, para esto se debe ir al botón **Import Dataset > From Excel**, en la ventana que aparece se deberá indicar el archivo, así como también la hoja, y si la primera fila cuenta como nombre de variable, entre otras cosas.



Import Excel Data

File/Url: C:/Users/JNavarro/Downloads/alfabet.xlsx

Data Preview:

PAIS	TASA
(character)	(character)
Afghanistan	38.168041229247997
Albania	97.553901672363295
Algeria	79.6083984375
American Samoa	..
Andorra	..
Angola	71.164161682128906
Antigua and Barbuda	..
Argentina	98.089973449707003

Previewing first 50 entries.

Import Options:

Name: alfabet ☒ First Row as Names

Sheet: Default NA: Default

Skip: 0 ☒ Open Data Viewer

Code Preview:

```
library(readxl)
alfabet <- read_excel("C:/Users/JNavarro/Downloads/alfabet.xlsx")
view(alfabet)
```

Import Cancel

La particularidad de **RStudio** es que ofrece la sintaxis relacionada al procedimiento que se ha hecho, al igual que en **Rcmdr**. La salida mostrada en la consola es la siguiente:

```
> library(readxl)
> alfabet = read_excel("C:/Users/JNavarro/Downloads/alfabet.xlsx")
> View(alfabet)
```

Importar datos de Stata

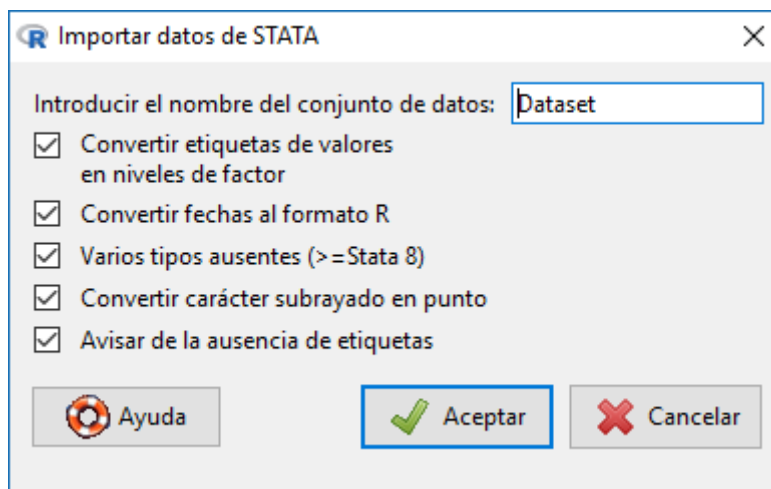
La forma directa de importar datos de Stata es utilizar el comando **read.dta()** del paquete **foreign**.

Rcmdr y RStudio ofrecen ambos, herramientas para la importación directa de datos con formato de Stata.

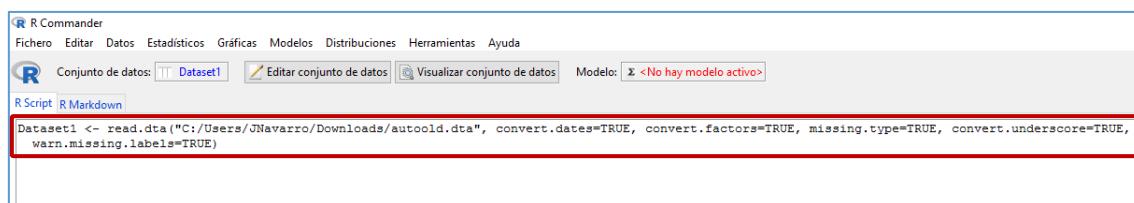
Rcmdr

Rcmdr permite importar datos de Stata, pero debe tenerse en cuenta de que la base de datos de Stata debe estar guardada en un formato de Stata antiguo (versión 12 o menor). El paquete usado es **foreign**, este paquete permite la importación de diversas varias de datos que no son de R propiamente.

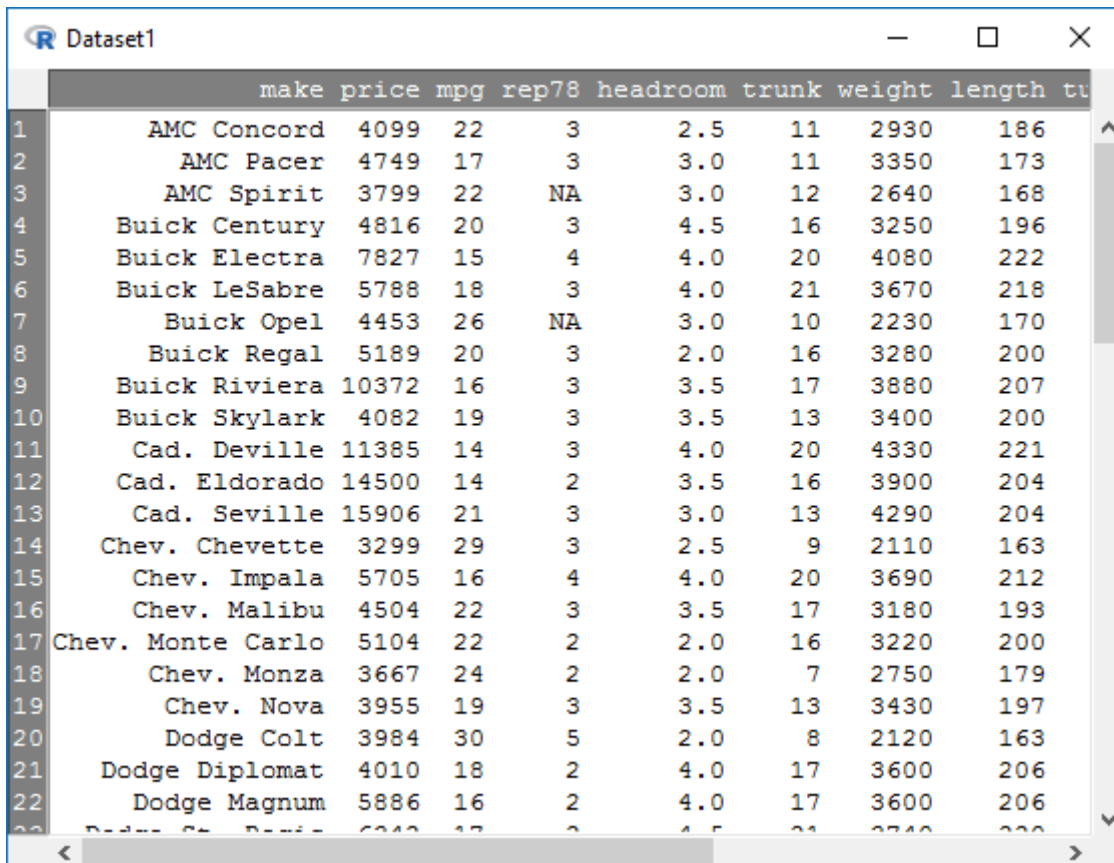
Para hacer la importación se debe ir a **Datos > Importar Datos > Desde Stata**, aparecerá una ventana en donde se deberá indicar características de la importación, como el nombre del archivo importado, entre otras cosas:



Una vez que se acepte se abrirá un explorador del sistema en donde se deberá indicar la ubicación del archivo, luego de eso se importará la base de datos, y la sintaxis del procedimiento se mostrará en la parte del RScript:



Para hacer la visualización de la base de datos se deberá dar clic en el botón de **Visualizar conjunto de datos** y los datos serán mostrados:



	make	price	mpg	rep78	headroom	trunk	weight	length	tu
1	AMC Concord	4099	22	3	2.5	11	2930	186	
2	AMC Pacer	4749	17	3	3.0	11	3350	173	
3	AMC Spirit	3799	22	NA	3.0	12	2640	168	
4	Buick Century	4816	20	3	4.5	16	3250	196	
5	Buick Electra	7827	15	4	4.0	20	4080	222	
6	Buick LeSabre	5788	18	3	4.0	21	3670	218	
7	Buick Opel	4453	26	NA	3.0	10	2230	170	
8	Buick Regal	5189	20	3	2.0	16	3280	200	
9	Buick Riviera	10372	16	3	3.5	17	3880	207	
10	Buick Skylark	4082	19	3	3.5	13	3400	200	
11	Cad. Deville	11385	14	3	4.0	20	4330	221	
12	Cad. Eldorado	14500	14	2	3.5	16	3900	204	
13	Cad. Seville	15906	21	3	3.0	13	4290	204	
14	Chev. Chevette	3299	29	3	2.5	9	2110	163	
15	Chev. Impala	5705	16	4	4.0	20	3690	212	
16	Chev. Malibu	4504	22	3	3.5	17	3180	193	
17	Chev. Monte Carlo	5104	22	2	2.0	16	3220	200	
18	Chev. Monza	3667	24	2	2.0	7	2750	179	
19	Chev. Nova	3955	19	3	3.5	13	3430	197	
20	Dodge Colt	3984	30	5	2.0	8	2120	163	
21	Dodge Diplomat	4010	18	2	4.0	17	3600	206	
22	Dodge Magnum	5886	16	2	4.0	17	3600	206	
23	Dodge St. Pierre	6343	17	2	4.5	21	3740	220	

RStudio

RStudio tiene herramientas para la importación de forma directa de un archivo de Stata.

Para importar un archivo de Stata en R se deberá ir a la sección en **Environment** en la interfaz del programa y elegir el botón de **Import Dataset**, se deberá seleccionar **From Stata**.

Aparecerá una ventana en donde se deberá indicar el archivo de Stata que se quiere importar:

Import Statistical Data

File/Url:
C:/Users/JNavarro/Downloads/auto.dta Browse...

Data Preview:

make Make and Model	price Price	mpg Mileage (mpg)	rep78 Repair Record 1978	headroom Headroom (in.)	trunk Trunk space (cu. ft.)	weight Weight (lbs.)	length Length (in.)	turn Turn Circle (ft.)	displacement Displacement (cu. in.)	gear_ratio Gear Ratio	foreign Car type
Plym. Champ	4425	34	5	2.5	11	1800	157	37	86	2.97	0
Plym. Horizon	4482	25	3	4.0	17	2200	165	36	105	3.37	0
Plym. Sapporo	6486	26	NaN	1.5	8	2520	182	38	119	3.54	0
Plym. Volare	4060	18	2	5.0	16	3330	201	44	225	3.23	0
Pont. Catalina	5798	18	4	4.0	20	3700	214	42	231	2.73	0
Pont. Firebird	4934	18	1	1.5	7	3470	198	42	231	3.08	0
Pont. Grand Prix	5222	19	3	2.0	16	3210	201	45	231	2.93	0
Pont. Le Mans	4723	19	3	3.5	17	3200	199	40	231	2.93	0

Previewing first 50 entries.

Import Options:

Name: Browse...

Model:

Format: ☒ Open Data Viewer

Code Preview:

```
library(haven)
auto <- read_dta("C:/Users/JNavarro/Downloads/auto.dta")
View(auto)
```

Import Cancel

La vista previa guiará el proceso de importación, y finalmente se mostrará en la consola de RStudio la salida de la importación con el paquete utilizado, que es el paquete **haven**, y el comando para el mismo:

```
> library(haven)
> auto = read_dta("C:/Users/JNavarro/Downloads/auto.dta")
> View(auto)
```

Exportar datos en R

R puede exportar a diversos formatos de archivos de datos, pero ni RStudio ni Rcmdr tienen opciones para exportar de forma directa, por lo que la exportación se hace usando los códigos.

Exportar a CSV

Para exportar a CSV debe hacerse uso del comando **write.csv()**. Este comando permitirá exportar los datos a formato CSV.

Suponiendo que se tiene abierta en R una base de datos que se llama **auto**, para que sea exportada en dicho formato se seguirá la siguiente sintaxis:

```
> write.csv(auto, 'auto.csv')
```

Esto indicará que se usará la base de datos **auto** para crear el archivo **auto.csv**, en el directorio de trabajo previamente indicado se encontrará la base de datos en el formato al cual se exportó.

Exportar a texto separado por tabulaciones

Se deberá usar el comando **write.table()**. La sintaxis basada en el ejemplo anterior es la siguiente:

```
> write.table(auto, 'auto.txt', sep='\t', dec='.')
```

Esto indicará que se está creando un archivo de formato **txt** llamado **auto**, que tiene como separadores de campos a tabuladores y como separador decimal a los puntos (.).

Exportar datos a Excel

Para exportar datos al formato de Stata se necesitará el paquete **WriteXLS**, el cual deberá ser instalado. El comando a utilizarse es **WriteXLS()**, la sintaxis del mismo es parecida a la de los anteriores:

```
> WriteXLS("auto", "auto.xls")
```

Esta sintaxis indicará que se creará un archivo de nombre **auto** basado en la data del mismo nombre.

Exportar datos a Stata

Para exportar datos a Stata se debe usar el comando **write.dta()** del paquete **foreign**, la sintaxis será la siguiente:

```
> write.dta(alfabet, 'alfabet.dta')
```

Esta sintaxis indicará que se creará un archivo de nombre **alfabet** basado en la data del mismo nombre que estará en formato **dta**.

Bibliografía

Fox, John. (2008). Iniciación a R Commander. Revisado el 21 de noviembre de 2016.

<http://www.uv.es/conesa/CursoR/material/Empezando-con-Rcmdr-es.pdf>

Recursos informáticos

RUG Barcelona - Importar - Exportar datos en R:

<https://rugbcn.wordpress.com/2012/10/08/963/>

Rcommander:

<http://www.rcommander.com/>

Universidad de Murcia - Entorno de trabajo R:

<http://www.um.es/ae/FEIR/10/#por-que-emplear-r>