

Importar bases de datos en formato Stata y SPSS en R.

Por: **Jeferson Ruiz** (nicaeda@gmail.com)

Hola amigos usuarios de R y cibernautas de nuestro blog amigo. En esta ocasión aprenderemos cómo importar datos en formato Stata y SPSS en R, bajo el entorno Windows. El primer paso es hacer doble clic en el ícono de R, si ya lo tenemos instalado.



R 1386 2.15.0

Una vez abierto el programa, R, nos muestra la consola:

Una captura de pantalla de la interfaz de usuario de R (RGui) en un entorno Windows. La ventana principal muestra un menú de opciones (Archivo, Editar, Visualizar, Misc, Paquetes, Ventanas, Ayuda) y una barra de herramientas con íconos para abrir, guardar, imprimir, etc. Abajo, una ventana titulada 'R Console' muestra el texto de bienvenida de R, incluyendo la versión (2.15.0), derechos de autor (2012 The R Foundation for Statistical Computing), ISBN (3-900051-07-0) y plataforma (1386-pc-mingw32/1386 (32-bit)). El texto continúa informando que R es un software libre sin garantía alguna y un proyecto colaborativo. Se muestran comandos como 'license()', 'contributors()', 'citation()', 'demo()', 'help()', 'help.start()' y 'q()' con sus respectivos usos. Al final, se indica '[Previously saved workspace restored]' y se muestra el prompt de comando '> |'.

Por lo que debemos de ejecutar el comando *library* (“foreign”), para que R nos pueda reconocer las bases de datos que vamos a importar. Pues la “library foreign” es el comando que R utiliza para reconocer las extensiones de distintos programas. Este comando reconoce las extensiones de programas como Stata, SPSS y Minitab, por mencionar algunos. Escribimos entonces el comando:

```
> library ("foreign")
```

Importar bases de datos en formato Stata y SPSS en R.

Por: Jeferson Ruiz (nicaeda@gmail.com)

>

Si lo hemos escrito bien, y si la pantalla no nos muestra un mensaje de error, quiere decir que hemos activado el paquete.

Ahora, el siguiente paso es importar una base de datos en formato Stata, utilizaremos una base de la encuesta de hogares de Nicaragua. El comando que R reconoce para importar bases de datos en formato Stata es "read.dta". Y la sintaxis es la siguiente:

Vamos a llamar a nuestra base de datos "stata", por lo que escribimos:

```
stata=read.dta
```

dado que R trabaja con un lenguaje de "objetos" primero ponemos nombre a la base de datos, que en este ejercicio la hemos llamado "stata", luego escribimos el símbolo "=" para continuar con el comando y dar la orden de importar los datos. He de decir que el símbolo es relativo, pues otros usuarios prefieren el símbolo: "<-". Sin embargo, en lo personal, yo prefiero el anterior para optimizar tiempo. Y porque me siento más cómodo. Usted lo puede hacer como prefiera.

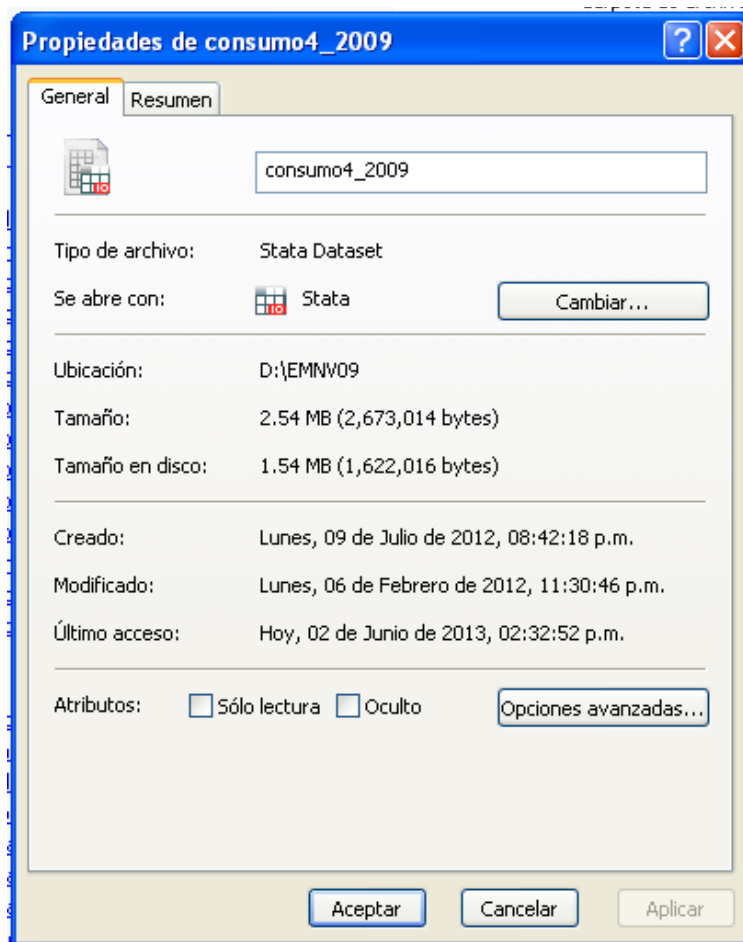
Bien, después de escribir; `stata=read.dta`, abrimos paréntesis y comillas: ("

Luego nos vamos donde tenemos el archivo y copiamos su dirección, haciendo clic derecho sobre el ícono de la base de datos y luego hacemos clic en "propiedades"

En mi computadora, tengo la base en una carpeta que se llama "EMNV09". Abro esa carpeta y me voy a la base que se llama "consumo4 2009". Hago clic derecho sobre ella y me voy a propiedades.

Importar bases de datos en formato Stata y SPSS en R.

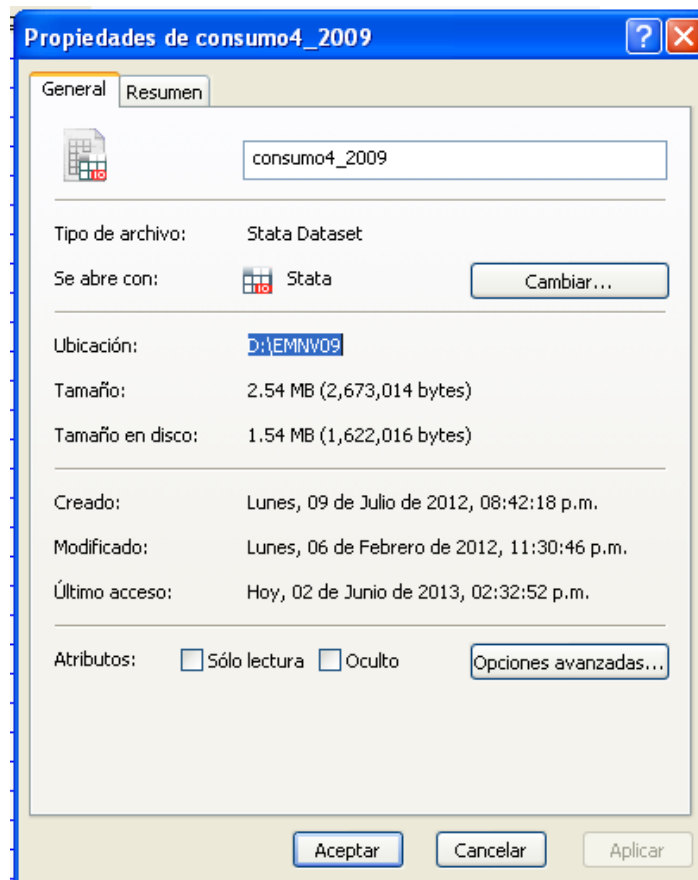
Por: Jeferson Ruiz (nicaeda@gmail.com)



Ahora sombro la ubicación y le doy copiar

Importar bases de datos en formato Stata y SPSS en R.

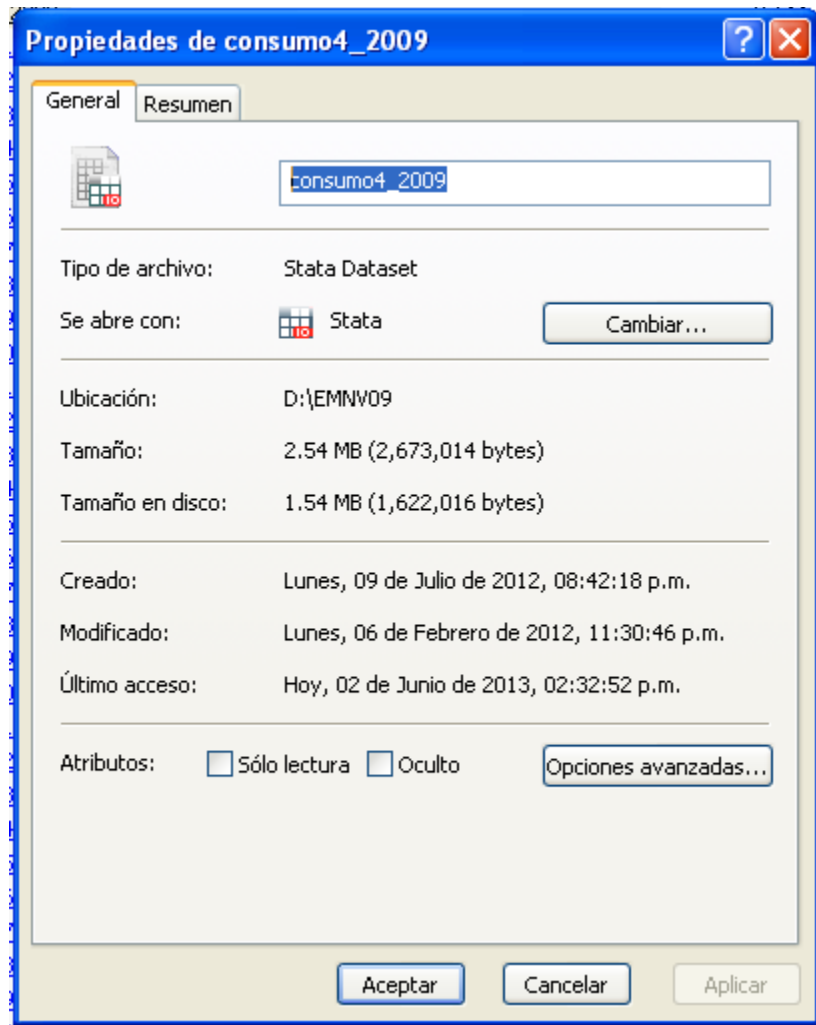
Por: Jeferson Ruiz (nicaeda@gmail.com)



Luego lo pego en la sintaxis: `stata=read.dta("D:\EMNV09`
Como notamos, la pleca hay que invertirla hasta quedar `"/`. Y, seguido de la dirección, volvemos a escribir una pleca puesto que hay que ir de nuevo a la ubicación: `stata=read.dta("D:/EMNV09/`
Hasta aquí, vamos construyendo la dirección para importar la base de datos. El siguiente paso es volver a la ubicación pero, ésta vez vamos a sombrear el nombre de la base:

Importar bases de datos en formato Stata y SPSS en R.

Por: Jeferson Ruiz (nicaeda@gmail.com)



Lo copiamos y lo pegamos después de la pleca, hasta quedar de esta manera: `stata=read.dta("D:/EMNV09/ consumo4_2009`

Hasta aquí, no hemos terminado, el último paso que damos es, escribir, seguido de el nombre de la base, la extensión ".dta" y, cerrar comillas y paréntesis, hasta quedar: `stata=read.dta("D:/EMNV09/consumo4_2009.dta")`.

Escribamos este comando en la consola de R:

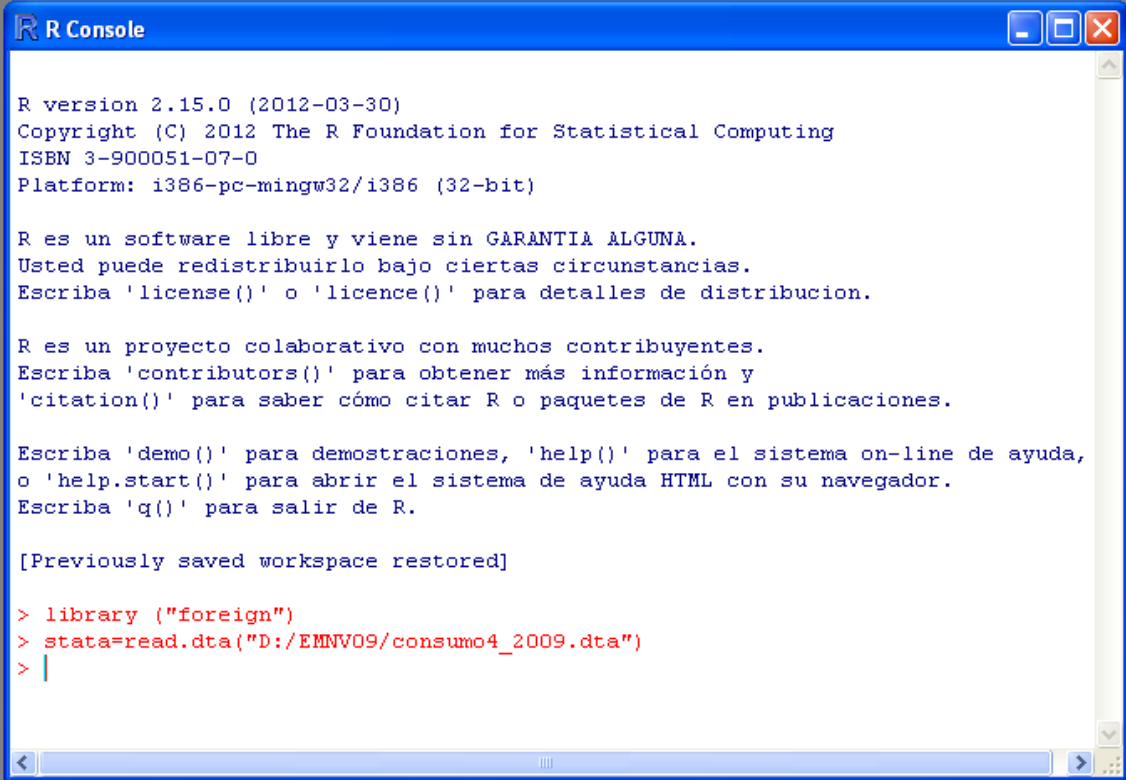
```
> stata=read.dta("D:/EMNV09/consumo4_2009.dta")
```

```
>
```

Si R no nos muestra error y, si aparece el símbolo ">", quiere decir que hemos tenido éxito.

Importar bases de datos en formato Stata y SPSS en R.

Por: Jeferson Ruiz (nicaeda@gmail.com)



```
R R Console

R version 2.15.0 (2012-03-30)
Copyright (C) 2012 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: i386-pc-mingw32/i386 (32-bit)

R es un software libre y viene sin GARANTIA ALGUNA.
Usted puede redistribuirlo bajo ciertas circunstancias.
Escriba 'license()' o 'licence()' para detalles de distribucion.

R es un proyecto colaborativo con muchos contribuyentes.
Escriba 'contributors()' para obtener más información y
'citation()' para saber cómo citar R o paquetes de R en publicaciones.

Escriba 'demo()' para demostraciones, 'help()' para el sistema on-line de ayuda,
o 'help.start()' para abrir el sistema de ayuda HTML con su navegador.
Escriba 'q()' para salir de R.

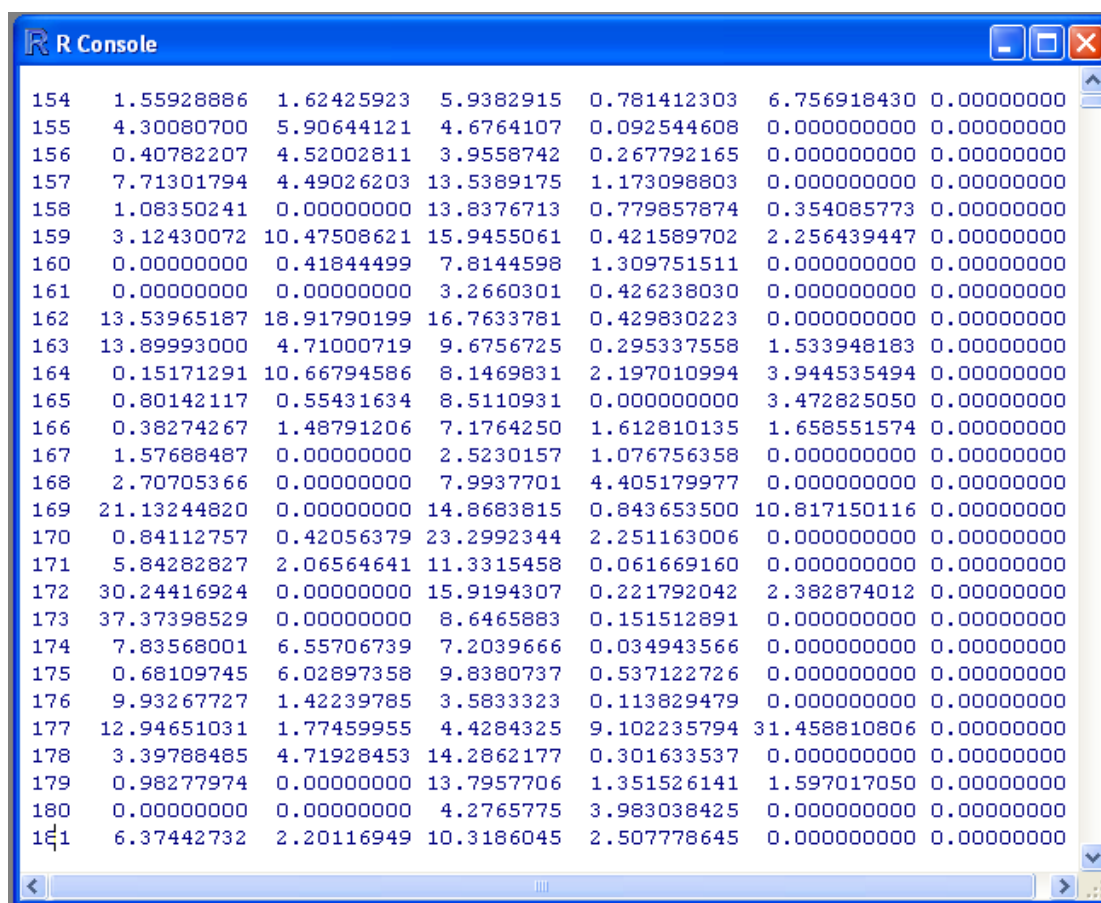
[Previously saved workspace restored]

> library ("foreign")
> stata=read.dta("D:/EMNV09/consumo4_2009.dta")
> |
```

Si queremos ver los datos, basta con solamente escribir el nombre que le dimos a la base de datos que, en este caso es stata y, la consola nos muestra los datos:

Importar bases de datos en formato Stata y SPSS en R.

Por: Jeferson Ruiz (nicaeda@gmail.com)



The screenshot shows the R Console window displaying a data table with 17 rows and 7 columns. The data is as follows:

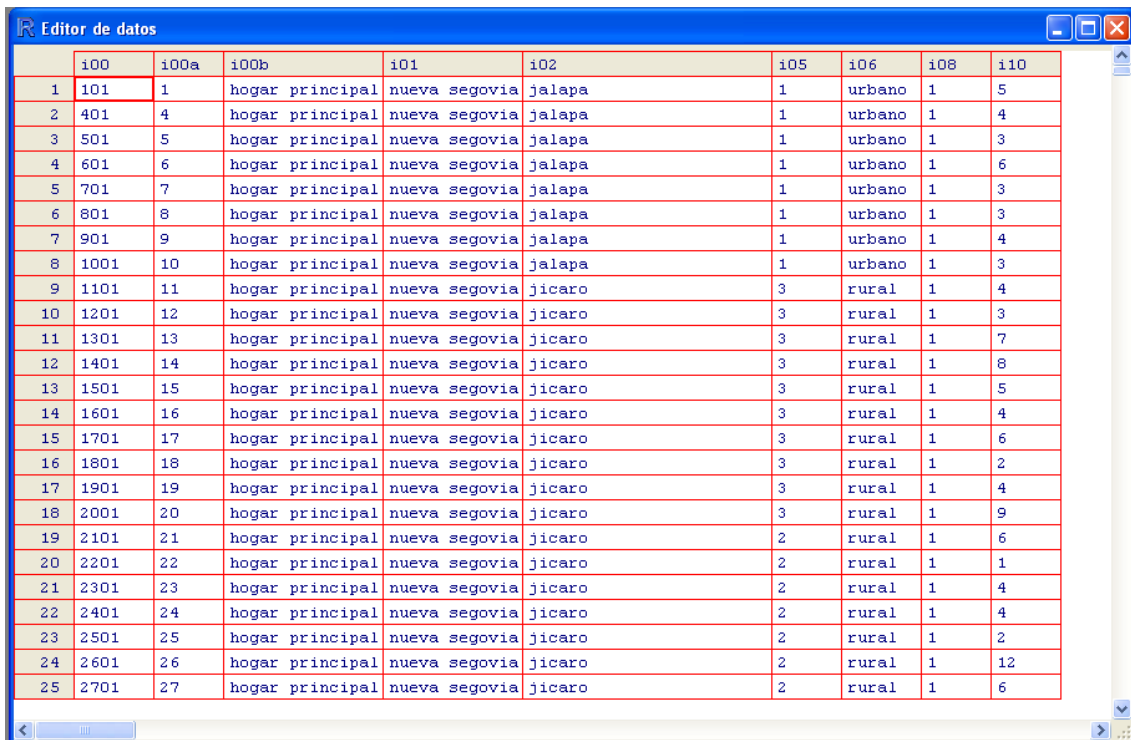
154	1.55928886	1.62425923	5.9382915	0.781412303	6.756918430	0.00000000
155	4.30080700	5.90644121	4.6764107	0.092544608	0.000000000	0.00000000
156	0.40782207	4.52002811	3.9558742	0.267792165	0.000000000	0.00000000
157	7.71301794	4.49026203	13.5389175	1.173098803	0.000000000	0.00000000
158	1.08350241	0.00000000	13.8376713	0.779857874	0.354085773	0.00000000
159	3.12430072	10.47508621	15.9455061	0.421589702	2.256439447	0.00000000
160	0.00000000	0.41844499	7.8144598	1.309751511	0.000000000	0.00000000
161	0.00000000	0.00000000	3.2660301	0.426238030	0.000000000	0.00000000
162	13.53965187	18.91790199	16.7633781	0.429830223	0.000000000	0.00000000
163	13.89993000	4.71000719	9.6756725	0.295337558	1.533948183	0.00000000
164	0.15171291	10.66794586	8.1469831	2.197010994	3.944535494	0.00000000
165	0.80142117	0.55431634	8.5110931	0.000000000	3.472825050	0.00000000
166	0.38274267	1.48791206	7.1764250	1.612810135	1.658551574	0.00000000
167	1.57688487	0.00000000	2.5230157	1.076756358	0.000000000	0.00000000
168	2.70705366	0.00000000	7.9937701	4.405179977	0.000000000	0.00000000
169	21.13244820	0.00000000	14.8683815	0.843653500	10.817150116	0.00000000
170	0.84112757	0.42056379	23.2992344	2.251163006	0.000000000	0.00000000
171	5.84282827	2.06564641	11.3315458	0.061669160	0.000000000	0.00000000
172	30.24416924	0.00000000	15.9194307	0.221792042	2.382874012	0.00000000
173	37.37398529	0.00000000	8.6465883	0.151512891	0.000000000	0.00000000
174	7.83568001	6.55706739	7.2039666	0.034943566	0.000000000	0.00000000
175	0.68109745	6.02897358	9.8380737	0.537122726	0.000000000	0.00000000
176	9.93267727	1.42239785	3.5833323	0.113829479	0.000000000	0.00000000
177	12.94651031	1.77459955	4.4284325	9.102235794	31.458810806	0.00000000
178	3.39788485	4.71928453	14.2862177	0.301633537	0.000000000	0.00000000
179	0.98277974	0.00000000	13.7957706	1.351526141	1.597017050	0.00000000
180	0.00000000	0.00000000	4.2765775	3.983038425	0.000000000	0.00000000
181	6.37442732	2.20116949	10.3186045	2.507778645	0.000000000	0.00000000

Otra manera de ver la base de datos completa, con sus encabezados y, si queremos modificar algún dato, solamente escribimos el comando:

> `fix(stata)`

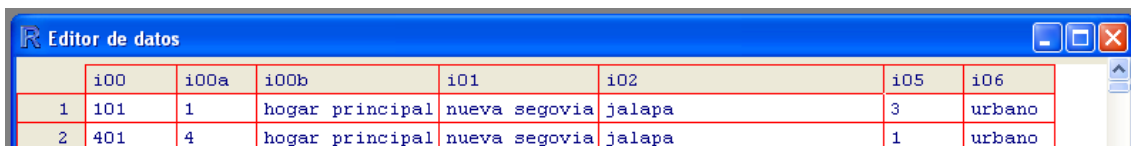
Importar bases de datos en formato Stata y SPSS en R.

Por: Jeferson Ruiz (nicaeda@gmail.com)



	i00	i00a	i00b	i01	i02	i05	i06	i08	i10
1	101	1	hogar principal	nueva segovia	jalapa	1	urbano	1	5
2	401	4	hogar principal	nueva segovia	jalapa	1	urbano	1	4
3	501	5	hogar principal	nueva segovia	jalapa	1	urbano	1	3
4	601	6	hogar principal	nueva segovia	jalapa	1	urbano	1	6
5	701	7	hogar principal	nueva segovia	jalapa	1	urbano	1	3
6	801	8	hogar principal	nueva segovia	jalapa	1	urbano	1	3
7	901	9	hogar principal	nueva segovia	jalapa	1	urbano	1	4
8	1001	10	hogar principal	nueva segovia	jalapa	1	urbano	1	3
9	1101	11	hogar principal	nueva segovia	jicaro	3	rural	1	4
10	1201	12	hogar principal	nueva segovia	jicaro	3	rural	1	3
11	1301	13	hogar principal	nueva segovia	jicaro	3	rural	1	7
12	1401	14	hogar principal	nueva segovia	jicaro	3	rural	1	8
13	1501	15	hogar principal	nueva segovia	jicaro	3	rural	1	5
14	1601	16	hogar principal	nueva segovia	jicaro	3	rural	1	4
15	1701	17	hogar principal	nueva segovia	jicaro	3	rural	1	6
16	1801	18	hogar principal	nueva segovia	jicaro	3	rural	1	2
17	1901	19	hogar principal	nueva segovia	jicaro	3	rural	1	4
18	2001	20	hogar principal	nueva segovia	jicaro	3	rural	1	9
19	2101	21	hogar principal	nueva segovia	jicaro	2	rural	1	6
20	2201	22	hogar principal	nueva segovia	jicaro	2	rural	1	1
21	2301	23	hogar principal	nueva segovia	jicaro	2	rural	1	4
22	2401	24	hogar principal	nueva segovia	jicaro	2	rural	1	4
23	2501	25	hogar principal	nueva segovia	jicaro	2	rural	1	2
24	2601	26	hogar principal	nueva segovia	jicaro	2	rural	1	12
25	2701	27	hogar principal	nueva segovia	jicaro	2	rural	1	6

Si por ejemplo en la variable i05, el primer número no es 1 y, tal vez es cero o cualquier otro número, basta con solamente hacer doble clic en la casilla y lo modificamos. Para ilustrar vamos a suponer que el primer número es 3.



	i00	i00a	i00b	i01	i02	i05	i06
1	101	1	hogar principal	nueva segovia	jalapa	3	urbano
2	401	4	hogar principal	nueva segovia	jalapa	1	urbano

Observe que lo modificamos sin ningún problema. Ahora lo único que queda es trabajar en nuestra base de datos.

Importar de SPSS.

Importar bases de datos en formato Stata y SPSS en R.

Por: Jeferson Ruiz (nicaeda@gmail.com)

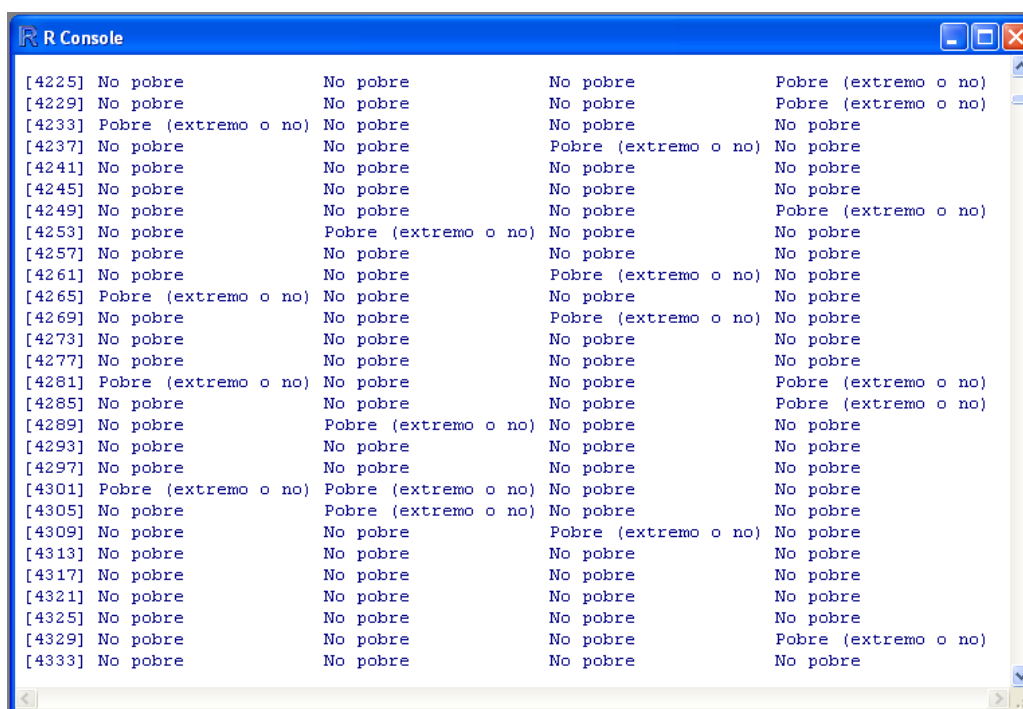
El procedimiento para importar bases de datos en formato SPSS, es prácticamente el mismo que hicimos con Stata, lo único que cambia es el comando para importar: "read.spss" y la extensión al final del nombre de la base de datos: ".sav"

Supongamos que esta vez vamos a trabajar con la encuesta "Datos de la vivienda y el hogar". A esta base de datos la vamos a llamar: "base". Escribimos pues el siguiente comando:

```
> base=read.spss("D:/EMNV09/DATOS ORIGINALES SAV/EMNV09-02 DATOS DE LA VIVIENDA Y EL HOGAR.sav")
```

```
>
```

Recuerde que, si queremos ver los datos, solamente escribimos el nombre que le dimos a la base de datos que, en este caso el nombre es: "base". Escribimos el nombre y damos "Enter".



```
R Console
[4225] No pobre          No pobre          No pobre          Pobre (extremo o no)
[4229] No pobre          No pobre          No pobre          Pobre (extremo o no)
[4233] Pobre (extremo o no) No pobre          No pobre          No pobre
[4237] No pobre          No pobre          Pobre (extremo o no) No pobre
[4241] No pobre          No pobre          No pobre          No pobre
[4245] No pobre          No pobre          No pobre          No pobre
[4249] No pobre          No pobre          No pobre          Pobre (extremo o no)
[4253] No pobre          Pobre (extremo o no) No pobre          No pobre
[4257] No pobre          No pobre          No pobre          No pobre
[4261] No pobre          No pobre          Pobre (extremo o no) No pobre
[4265] Pobre (extremo o no) No pobre          No pobre          No pobre
[4269] No pobre          No pobre          Pobre (extremo o no) No pobre
[4273] No pobre          No pobre          No pobre          No pobre
[4277] No pobre          No pobre          No pobre          No pobre
[4281] Pobre (extremo o no) No pobre          No pobre          Pobre (extremo o no)
[4285] No pobre          No pobre          No pobre          Pobre (extremo o no)
[4289] No pobre          Pobre (extremo o no) No pobre          No pobre
[4293] No pobre          No pobre          No pobre          No pobre
[4297] No pobre          No pobre          No pobre          No pobre
[4301] Pobre (extremo o no) Pobre (extremo o no) No pobre          No pobre
[4305] No pobre          Pobre (extremo o no) No pobre          No pobre
[4309] No pobre          No pobre          Pobre (extremo o no) No pobre
[4313] No pobre          No pobre          No pobre          No pobre
[4317] No pobre          No pobre          No pobre          No pobre
[4321] No pobre          No pobre          No pobre          No pobre
[4325] No pobre          No pobre          No pobre          No pobre
[4329] No pobre          No pobre          No pobre          Pobre (extremo o no)
[4333] No pobre          No pobre          No pobre          No pobre
```

Igual, recuerde que podemos modificar los datos con el comando "fix".

Con esto hemos terminado. En el siguiente tutorial, aprenderemos a copiar y pegar desde una hoja de Excel a R.

Recuerden que, cualquier pregunta, duda o sugerencia pueden escribir a:

Jeferson Ruiz (nicaeda@gmail.com)
Deybi Morales (morales.economia@gmail.com)
Visiten nuestro blog: <http://aulaweb.uca.edu.ni/blogs/rnica>
Fundadores del Grupo de usuarios R Nicaragua