

Introducción a Stata

Parte 2

Diego Salmerón Martínez

CIBER de Epidemiología y Salud Pública

Consejería de Sanidad de Murcia

Instituto Murciano de Investigación Biosanitaria – Arrixaca

Universidad de Murcia

- Crear y modificar variables
- Trabajar con fechas
- Selección de casos y/o variables
- Juntar bases de datos
- Exportar resultados
- Guardar y exportar bases de datos

Crear y modificar variables

```
generate newvar = exp
```

```
recode variable_discreta regla
```

```
replace oldvar = exp
```

Se pueden utilizar con la opción **if**

Se aprende haciendo !



sample_Breast_.xlsx

Casos de cáncer de mama, 1995-2007, seguimiento hasta 31/12/2011

Fechas: nacimiento, incidencia, seguimiento

Estado vital, topografía y morfología, sexo, Identificación del tumor

```
import excel sample_Breast_.xlsx, sheet("Sheet1") firstrow
```

Generar la edad al diagnóstico en días, a partir de las fechas de nacimiento y diagnóstico

Generar la edad en años y en años cumplidos

Generar grupos de edad: **recode, replace + if**

Trabajar con fechas

Formatos

`format diagmdy %d`

Editor de Datos (Edición) - [Sin_título]

Archivo Edición Ver Datos Herramientas

diagmdy[1] 15nov1996

	diagmdy	vsmdy	topo	morfo	nacmdy
1	11/15/1996	3/29/2003	C504	8520	7/10/1937
2	2/13/2004	7/21/2006	C505	8503	2/12/1930
3	2/15/2001	12/31/2011	C509	8500	10/25/1925
4	5/12/2004	3/22/2005	C509	8500	6/8/1933
5	11/8/2000	12/31/2011	C509	9999	6/3/1958
6	10/14/2002	12/31/2011	C509	8500	6/20/1969
7	9/25/2007	4/14/2011	C509	8500	10/11/1940
8	1/1/1995	12/31/2011	C508	8500	5/4/1942
9	10/5/2007	12/31/2011	C509	8500	11/22/1946
10	2/9/1999	11/20/2009	C509	8500	6/19/1960
11	4/23/1999	6/17/1999	C509	8500	9/6/1915
12	11/23/2005	11/30/2006	C509	9999	6/9/1921
13	6/27/1997	12/31/2011	C509	8481	8/1/1947
14	9/12/2003	6/3/2009	C504	8500	7/28/1939
15	4/6/2000	12/31/2011	C504	8500	4/4/1954
16	5/19/1998	2/20/2007	C504	8500	12/29/1949

También podemos cambiar el formato en la ventana Propiedades

Editor de Datos (Edición) - [Sin_título]

Archivo Edición Ver Datos Herramientas

1C 15nov1996 DMY

	diagmdy	vsmdy	topo	morfo	nacmdy	dead	sex	id	agediagind~s	edad_dx	edad5c	agegroup
1	15nov1996	3/29/2003	C504	8520	Jul 10, 1937	1	2	1	21678	59.35113	3	3
2	13feb2004	7/21/2006	C505	8503	Feb 12, 1930	1	2	2	27029	74.00137	4	4
3	15feb2001	12/31/2011	C509	8500	Oct 25, 1925	0	2	3	27507	75.31006	5	5
4	12may2004	3/22/2005	C509	8500	Jun 08, 1933	1	2	4	25905	70.92403	4	4
5	08nov2000	12/31/2011	C509	9999	Jun 03, 1958	0	2	5	15498	42.43121	1	1
6	14oct2002	12/31/2011	C509	8500	Jun 20, 1969	0	2	6	12165	33.31691	1	1
7	25sep2007	4/14/2011	C509	8500	Oct 11, 1940	1	2	7	24455	60.95414	4	4
8	01jan1995	12/31/2011	C508	8500	May 04, 1942	0	2	8	19235	52.66205	2	2
9	05oct2007	12/31/2011	C509	8500	Nov 22, 1946	0	2	9	22232	60.8679	3	3
10	09feb1999	11/20/2009	C509	8500	Jun 19, 1960	1	2	10	14113	38.63929	1	1
11	23apr1999	6/17/1999	C509	8500	Sep 06, 1915	1	2	11	30544	83.62492	5	5
12	23nov2005	11/30/2006	C509	9999	Jun 09, 1921	1	2	12	30848	84.45722	5	5
13	27jun1997	12/31/2011	C509	8481	Aug 01, 1947	0	2	13	18228	49.90554	2	2
14	12sep2003	6/3/2009	C504	8500	Jul 28, 1939	1	2	14	23421	64.12321	3	3
15	06apr2000	12/31/2011	C504	8500	Apr 04, 1954	0	2	15	16803	46.00411	2	2
16	19may1998	2/20/2007	C504	8500	Dec 29, 1949	1	2	16	17673	48.38604	2	2
17	01jan1999	12/31/2011	C504	8500	Oct 19, 1925	0	2	17	26736	73.19918	4	4
18	25jul1996	8/25/2003	C509	8520	Jan 01, 1928	1	2	18	25043	68.564	4	4
19	03mar1999	10/18/2002	C503	8500	Jul 15, 1927	1	2	19	26164	71.63313	4	4
20	01nov1995	10/22/2005	C508	8500	Feb 11, 1932	1	2	20	23274	63.72074	3	3
21	04dec1996	12/31/2011	C509	8500	Apr 23, 1941	0	2	21	20313	55.61396	3	3

Variables

Filtrar variables aquí

<input checked="" type="checkbox"/>	Nombre	Etiqueta
<input checked="" type="checkbox"/>	diagmdy	diagmdy
<input checked="" type="checkbox"/>	vsmdy	vsmdy
<input checked="" type="checkbox"/>	topo	topo
<input checked="" type="checkbox"/>	morfo	morfo
<input checked="" type="checkbox"/>	nacmdy	nacmdy
<input checked="" type="checkbox"/>	dead	dead
<input checked="" type="checkbox"/>	sex	sex
<input checked="" type="checkbox"/>	id	id
<input checked="" type="checkbox"/>	agediagindays	
<input checked="" type="checkbox"/>	edad_dx	
<input checked="" type="checkbox"/>	edad5c	RECODE of edad_dx
<input checked="" type="checkbox"/>	agegroup	

Variables Copia temporal

Propiedades

Variables	
Nombre	diagmdy
Etiqueta	diagmdy
Typo	int
Formato	%d ...
Etiqueta de valor	
Notas	

Fechas y sus componentes

Sacar el día, mes y año de la fecha:

`day, month, year`

Construir la fecha desde el día, el mes y el año:

`mdy`

Se aprende haciendo !



Obtener el día, el mes y el año de la fecha del diagnóstico

Recomponer la fecha del diagnóstico en una nueva variable

Agrupar el año de diagnóstico en tres periodos

Seleccionar casos y/o variables

keep/drop if exp

Ejemplo: Seleccionamos el periodo 1

keep/drop variables

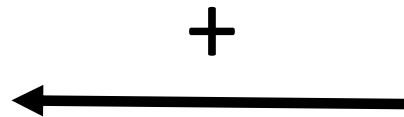
Ejemplo: Nos quedamos con las variables de fechas de incidencia y seguimiento, la edad categórica, el periodo, el identificador del tumor y el estado vital

Juntar bases de datos

Añadir variables de Datos2 a Datos1, mediante variables clave

```
use Datos1
```

```
merge var using Datos2
```



Variables clave

Información para añadir

```
merge id using datos2.dta, sort
```

(note: you are using old merge syntax; see [D] merge for new syntax)

```
tab _merge
```

_merge	Freq.	Percent	Cum.
2	735	27.67	27.67
3	1,921	72.33	100.00
Total	2,656	100.00	

```
keep if _merge==3
```

(735 observations deleted)

```
drop _merge
```

1	master	observation appeared in master only
2	using	observation appeared in using only
3	match	observation appeared in both
4	match_update	observation appeared in both, missing values updated
5	match_conflict	observation appeared in both, conflicting nonmissing values



diagmdy[1]

15nov1996

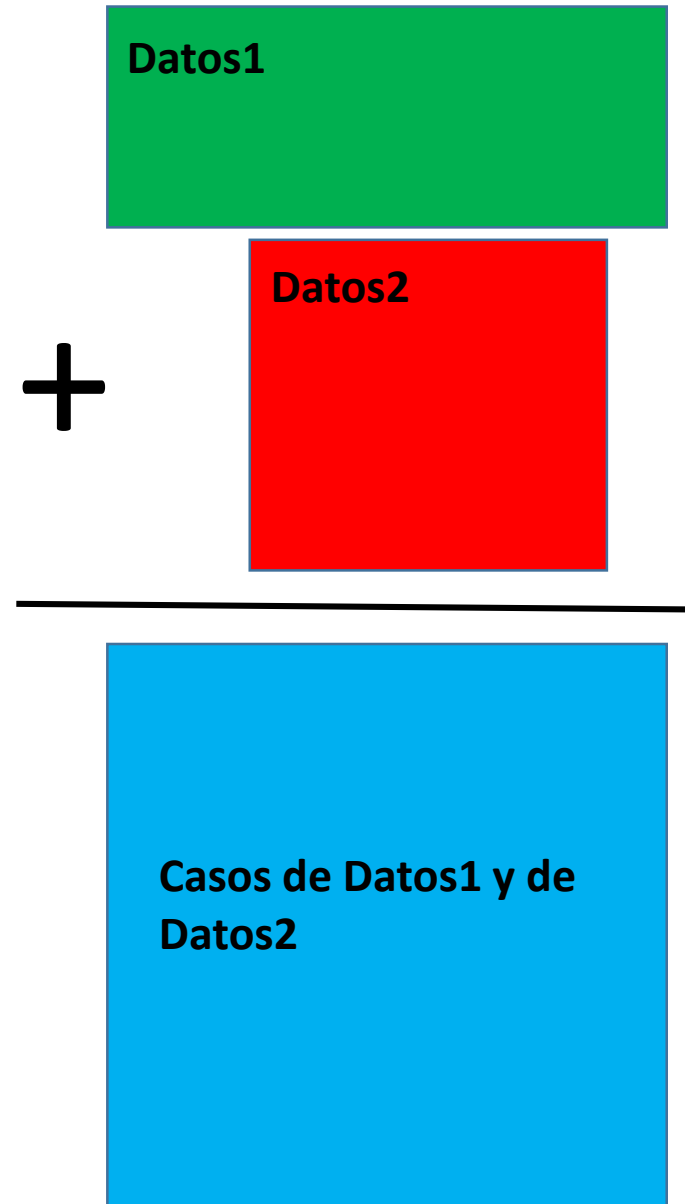
	diagmdy	vsmdy	topo	morfo	nacmdy	dead	sex	id	multiples	lateralidad
1	11/15/1996	3/29/2003	C504	8520	7/10/1937	1	2	1	1	1
2	2/13/2004	7/21/2006	C505	8503	2/12/1930	1	2	2	0	3
3	2/15/2001	12/31/2011	C509	8500	10/25/1925	0	2	3	0	1
4	5/12/2004	3/22/2005	C509	8500	6/8/1933	1	2	4	0	2
5	11/8/2000	12/31/2011	C509	9999	6/3/1958	0	2	5	0	1
6	10/14/2002	12/31/2011	C509	8500	6/20/1969	0	2	6	0	2
7	9/25/2007	4/14/2011	C509	8500	10/11/1940	1	2	7	0	1
8	1/1/1995	12/31/2011	C508	8500	5/4/1942	0	2	8	0	3
9	10/5/2007	12/31/2011	C509	8500	11/22/1946	0	2	9	1	1
10	2/9/1999	11/20/2009	C509	8500	6/19/1960	1	2	10	1	1
11	4/23/1999	6/17/1999	C509	8500	9/6/1915	1	2	11	0	1
12	11/23/2005	11/30/2006	C509	9999	6/9/1921	1	2	12	0	2
13	6/27/1997	12/31/2011	C509	8481	8/1/1947	0	2	13	0	2
14	9/12/2003	6/3/2009	C504	8500	7/28/1939	1	2	14	1	3
15	4/6/2000	12/31/2011	C504	8500	4/4/1954	0	2	15	0	2
16	5/19/1998	2/20/2007	C504	8500	12/29/1949	1	2	16	1	3
17	1/1/1999	12/31/2011	C504	8500	10/19/1925	0	2	17	0	2
18	7/25/1996	8/25/2003	C509	8520	1/1/1928	1	2	18	0	1
19	3/3/1999	10/18/2002	C503	8500	7/15/1927	1	2	19	1	3
20	11/1/1995	10/22/2005	C508	8500	2/11/1932	1	2	20	1	3
21	12/4/1996	12/31/2011	C509	8500	4/23/1941	0	2	21	1	3
22	3/28/2003	12/31/2011	C509	8500	4/19/1963	0	2	22	0	2

Juntar bases de datos

Añadir casos

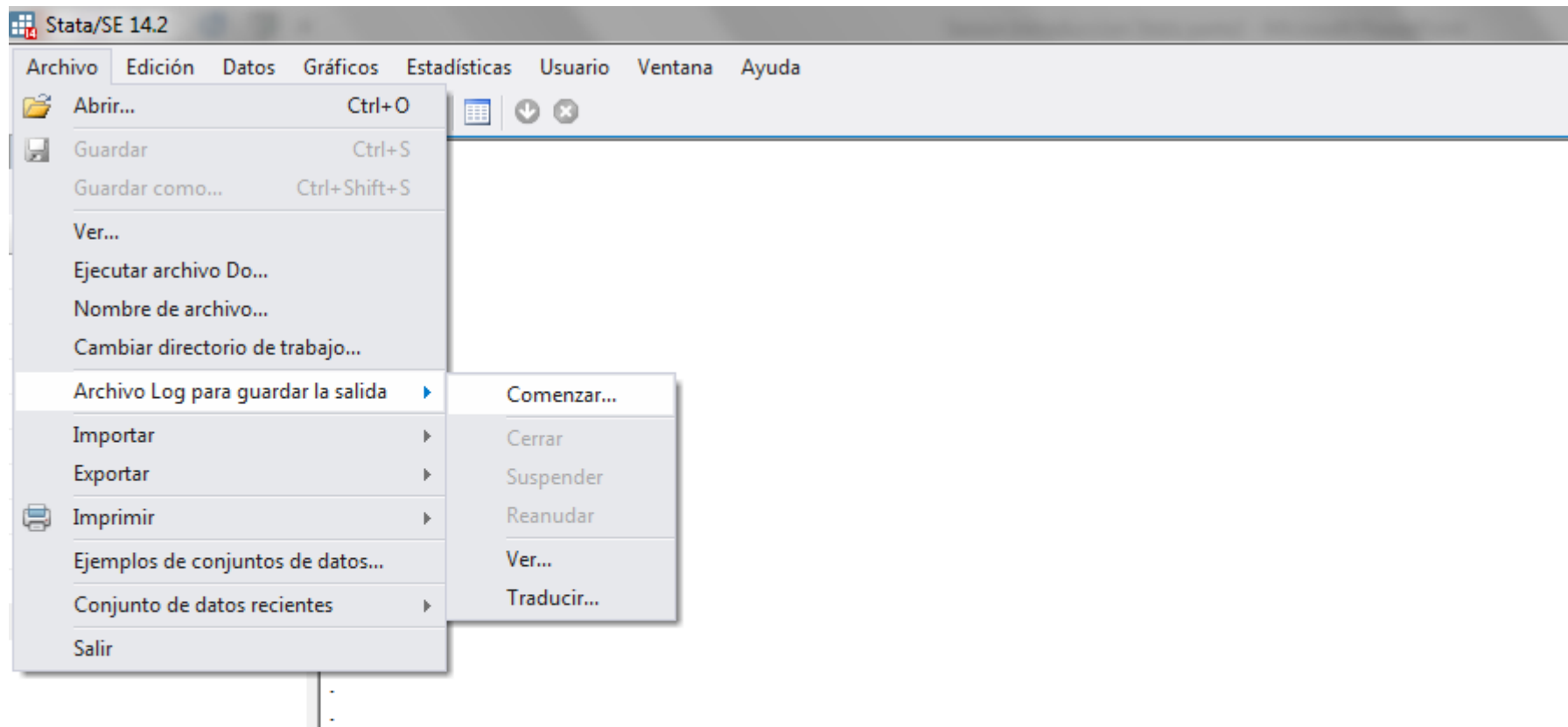
```
use Datos1
```

```
append using Datos2
```

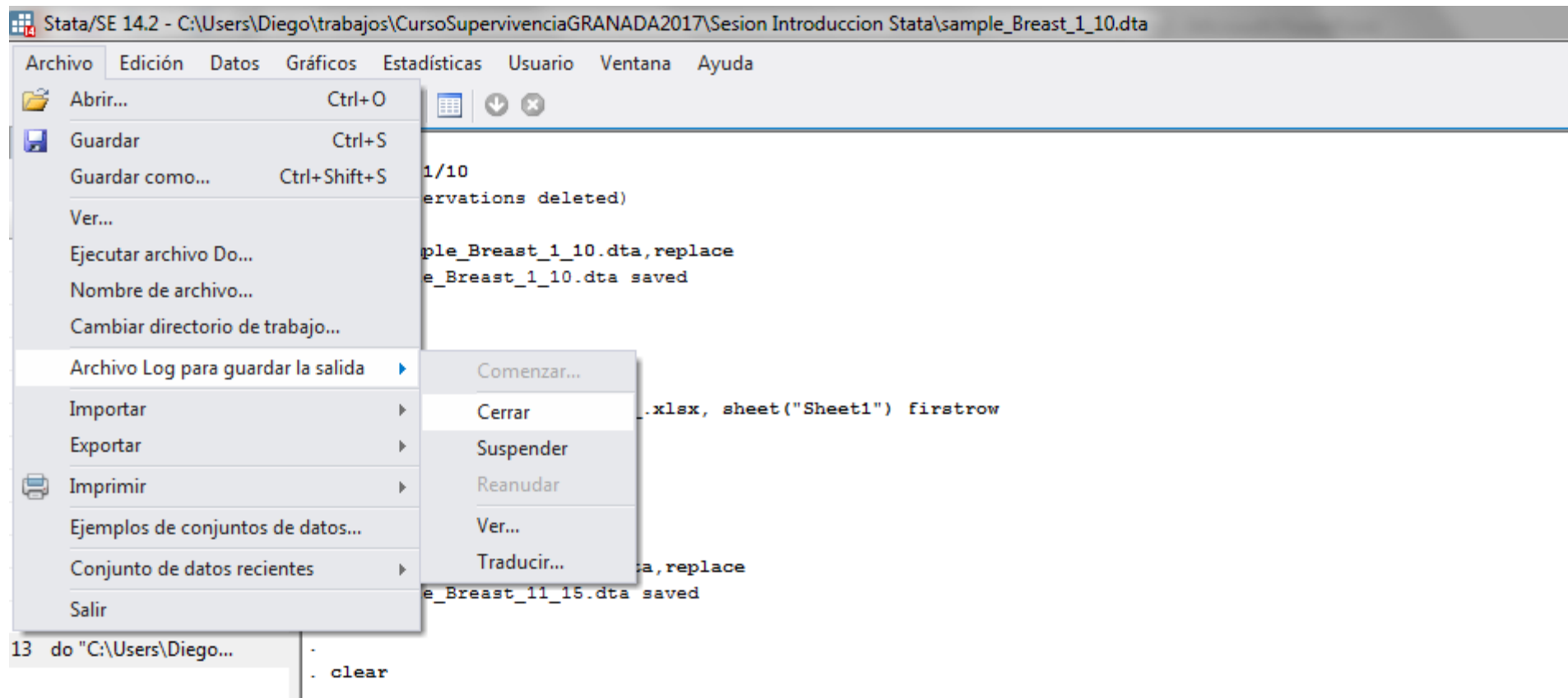


Guardar trabajo realizado: log

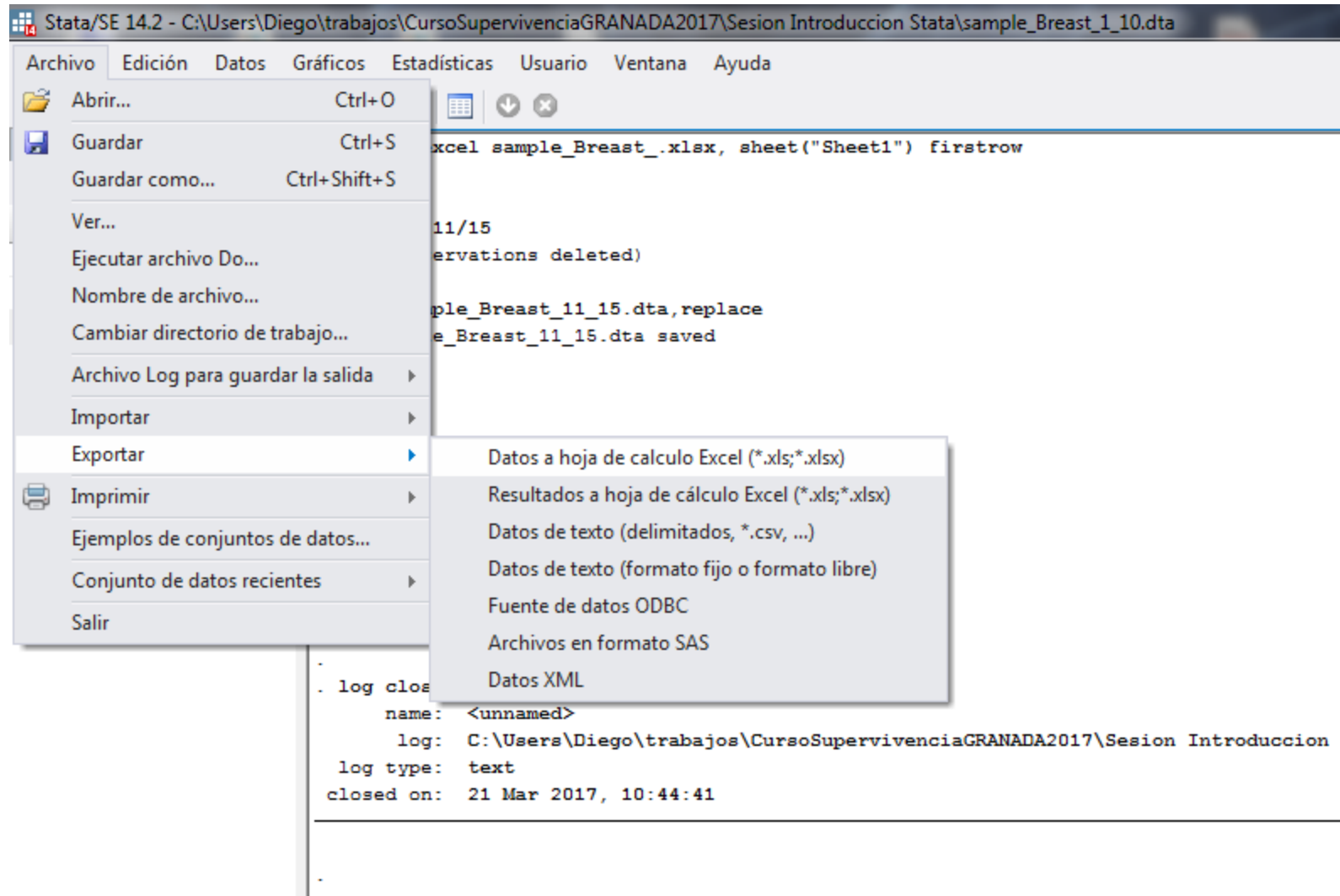
`log using "fichero.log"`



log close



Guardar y exportar bases de datos



Gracias por vuestra atención