

STATA Básico

Interfaz y comandos

Renzo Álvarez Carcheri

CENTEC

renzo.alvarez.carcheri@gmail.com

renzoac93@hotmail.com

September 19, 2021



- 1 Introducción
- 2 Interfaz
- 3 Comandos
- 4 Principales comandos I
- 5 Introducción a la programación

Conceptos

- Stata es un software estadístico que permite, entre otras funciones, la manipulación y visualización de datos, el análisis estadístico y reproducción de informes.
- Cuenta con una gran diversidad de procedimientos estadísticos que pueden ser aplicados en diferentes áreas y sectores, por lo que es ampliamente utilizado en investigación académica así como en entidades gubernamentales, financieras, comerciales y de servicios.
- Se considera uno de los software estadísticos más completos y más fáciles de usar, puesto que cuenta con una agradable interfaz gráfica que se puede utilizar a través de ventanas o con una sintaxis de comandos intuitiva y potente.

Renzo Álvarez Carcheri



- 1 Introducción
- 2 Interfaz**
- 3 Comandos
- 4 Principales comandos I
- 5 Introducción a la programación

Entorno de trabajo

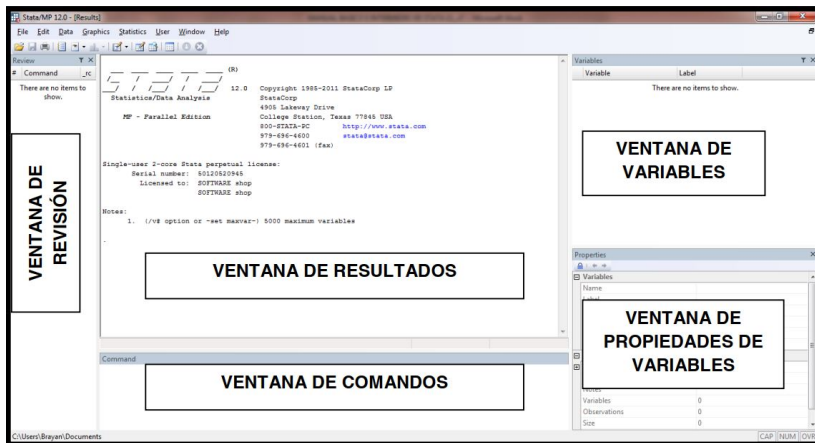


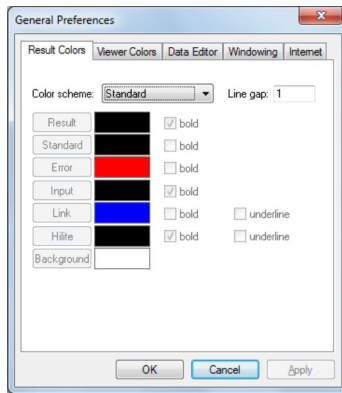
Figure 2: Entorno STATA

Ventanas de STATA

1. **Ventana de revisión:** Muestra los comandos que se han utilizado durante la sesión. Solo los resultados más recientes son visibles en esta pantalla.
2. **Ventana de comandos:** Sirve para manejar STATA de forma interactiva; es decir, se emplea para crear las líneas de comandos y llevar a cabo las aplicaciones disponibles en el software.
3. **Ventana de variables:** Informa sobre las variables que están disponibles en nuestra base de datos para realizar las diversas aplicaciones.
4. **Ventana de resultados:** Permite visualizar los resultados de los estadísticos que pedimos calcular o de los modelos que solicitamos estimar.
5. **Ventana de propiedades:** Permite editar las propiedades de las variables y de la base de datos, así como ver las propiedades generales.

Preferencias

- No obstante, el tamaño, posición y diseño de ventanas puede cambiarse con la opción: **Editar** → **Preferencias**.



Entorno Stata

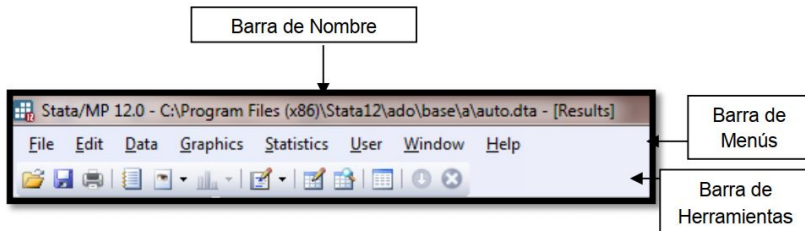


Figure 3: Menú de STATA

Menú de ayuda

El menú de ayuda permite:

- Obtener ayuda sobre algún comando de Stata.
- Listar y descargar las últimas actualizaciones del programa.
- Acceder a lugares de interés en el sitio Web de Stata.
- Instalar programas de Stata escritos por otros usuarios.

Menú de ayuda

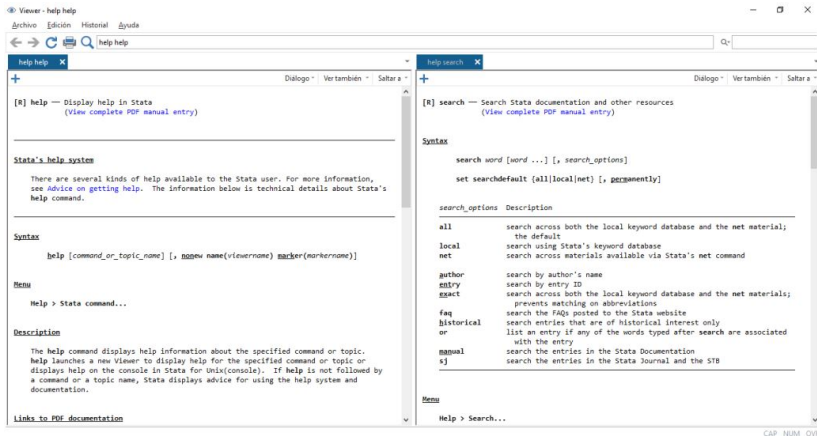


Figure 4: Ayuda de Stata (*Help*)

Stata editor

La ventana **Stata editor** permite navegar y modificar los datos como si fuese una hoja Excel. Mientras que **Stata editor browser** permite visualizar datos, más no modificarlos.

The screenshot shows the Stata Data Editor window titled "Data Editor (Edit) - [auto]". The main area displays a dataset of cars with the following columns: make, price, mpg, rep78, headroom, trunk, weight, length, turn, and displacement. The data is sorted by make. The right sidebar contains a "Variables" panel with a list of variables and their labels, and a "Properties" panel showing details for the selected variable "make".

make	price	mpg	rep78	headroom	trunk	weight	length	turn	displacement
1 AMC Concord	4,099	22	3	2.5	11	2,930	186	40	121
2 AMC Pacer	4,749	17	3	3.0	11	3,350	173	40	258
3 AMC Spirit	3,799	22	.	3.0	12	2,640	168	35	121
4 Buick Century	4,816	20	3	4.5	16	3,250	196	40	196
5 Buick Electra	7,827	15	4	4.0	20	4,080	222	43	350
6 Buick LeSabre	5,788	18	3	4.0	21	3,670	218	43	231
7 Buick Opel	4,453	26	.	3.0	10	2,230	170	34	304
8 Buick Regal	5,189	20	3	2.0	16	3,280	200	42	196
9 Buick Riviera	10,372	16	3	3.5	17	3,880	207	43	231
10 Buick Skylark	4,082	19	3	3.5	13	3,400	200	42	231
11 Cad. Deville	11,385	14	3	4.0	20	4,330	221	44	425
12 Cad. Eldorado	14,500	14	2	3.5	16	3,900	204	43	350
13 Cad. Seville	15,906	21	3	3.0	13	4,290	204	45	350
14 Chev. Chevette	3,299	29	3	2.5	9	2,110	163	34	231
15 Chev. Impala	5,705	16	4	4.0	20	3,690	212	43	250
16 Chev. Malibu	4,504	22	3	3.5	17	3,180	193	31	200
17 Chev. Monte Carlo	5,104	22	2	2.0	16	3,220	200	41	200
18 Chev. Monza	3,667	24	2	2.0	7	2,750	179	40	151
19 Chev. Nova	3,955	19	3	3.5	13	3,430	197	43	250
20 Dodge Colt	3,984	30	5	2.0	8	2,120	163	35	98
21 Dodge Diplomat	4,010	18	2	4.0	17	3,600	206	46	318

Variables

Variable	Label
<input checked="" type="checkbox"/> make	Make and Model
<input checked="" type="checkbox"/> price	Price
<input checked="" type="checkbox"/> mpg	Mileage (mpg)
<input checked="" type="checkbox"/> rep78	Repair Record 1978
<input checked="" type="checkbox"/> headroom	Headroom (in.)
<input checked="" type="checkbox"/> trunk	Trunk space (cu. ft.)
<input checked="" type="checkbox"/> weight	Weight (lbs.)
<input checked="" type="checkbox"/> length	Length (in.)
<input checked="" type="checkbox"/> turn	Turn Circle (ft.)

Properties

Variables

Name	make
Label	Make and Model
Type	str18
Format	%-18s
Value Label	
Notes	

Data

Filename	auto.dta
Label	1978 Automobile

Ready Vars: 12 Order: Dataset Obs: 74 Filter: Off Mode: Edit CAP NUM

Tipos de archivo

Tipo de Archivo	Extensión
Archivos de datos	*.dta
Archivos gráficos	*.gph
Bitácoras de salida	*.smcl
Archivos de comandos	*.do
Archivos de programación	*.ado

Figure 6: Tipos de archivo

1 Introducción

2 Interfaz

3 Comandos

Estructura de comandos

4 Principales comandos I

5 Introducción a la programación

1 Introducción

2 Interfaz

3 Comandos

Estructura de comandos

4 Principales comandos I

5 Introducción a la programación

Estructura de comandos

- En el menú de ayuda *Help* podemos encontrar información sobre todos los comandos que ofrece el STATA.
- Utilizaremos la ventana de comandos para ejecutar dichos comandos.
- La estructura de los comandos es la que sigue:



- Lo que va en *corchetes* es opcional.

Estructura de comandos

Se pueden observar:

- **[varlist]**: Es la lista de variables.
- **[if]**: Debe estar seguida de una expresión lógica para que solo los datos que cumplan con la condición puedan incluirse en dicho análisis.
- **[in]**: Permite indicar el rango de los datos que se desean analizar.
- **[weight]**: Son los pesos o las ponderaciones.
- **[options]**: Son las opciones de cada comando que se está utilizando. Siempre se antepone una coma a las opciones.

Importando datos

Datos con formato ".dta"

- Es el formato propio de Stata. Se usa el comando "use" para abrir estos archivos.

Datos con formatos distintos al ".dta"

- Es usual trabajar con archivos Excel (.xlsx y .xls).
- Los archivos almacenados como texto y separados por algún carácter (.csv).
- También se permite el uso de formatos menos comunes como ODBC, SAS, etc.

Importando datos internos

- Una base de datos interna de Stata se llama *auto*, la cual se puede abrir con el comando **sysuse**.
- En primer lugar, se debe indicar la ruta del directorio de trabajo con el comando **cd** (*current directory*).

```
cd "C:/Users/renzo/OneDrive/Curso STATA/clase1"
```

- Una vez definido el directorio de trabajo, se guardará la base de datos con los cambios realizados con el comando:

```
save auto.dta
```

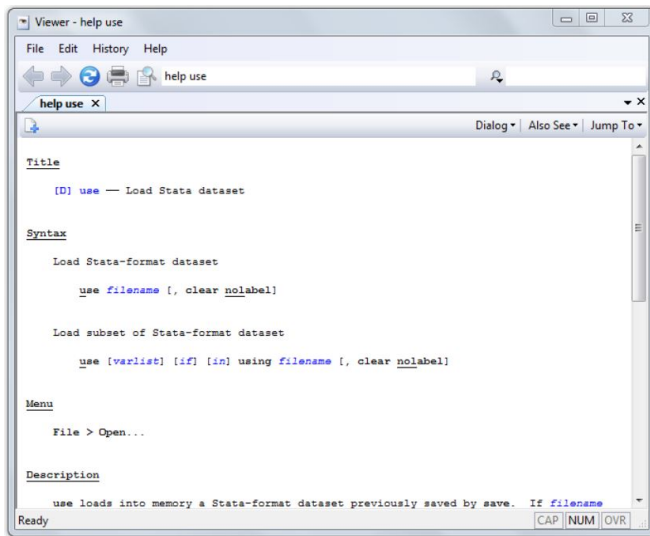
- Finalmente, se puede cargar la base de datos guardada usando:

```
use auto.dta
```

- Siempre puede encontrar información adicional del comando **help** seguido del comando que desea utilizar.

```
help use
```

Importando datos



Importando datos externos

Archivos Excel

- La sintaxis a utilizar es la siguiente:

```
import excel "$main/Datos.xlsx", sheet("Anuales") firstrow clear
```

Archivos .csv

- A diferencia del anterior, el comando requiere especificar el carácter que esta siendo utilizado como separador.

```
import delimited "$data/Data_Carros.csv", delimiter(comma) varnames(4) clear
```

- 1 Introducción
- 2 Interfaz
- 3 Comandos
- 4 Principales comandos I**
- 5 Introducción a la programación

Principales comandos I

Principales comandos 1	
help	Ayuda de programa STATA para comandos
use	Abrir una base de datos en STATA (extensión .dta)
codebook	Muestra variables con estadísticos descriptivos
describe	Describe contenido de datos en memoria o sobre el disco
save	Guarda los datos en disco con extensión *.dta
list	Lista de valores de variable
drop	Borrar variable o dato en memoria
keep	Borra variables o datos no seleccionados

Conociendo la data

- Por un lado, el comando **describe** le permite describir el contenido de datos en la memoria del computador.

```
describe make rice mpg
```

- Por otro lado, el comando **codebook** brinda información más detallada de las variables.

```
codebook price
```

- El comando **summarize** brinda reportes de los principales estadísticos de la base de datos.

```
summarize price mpg rep78 trunk weight length
```


Operadores

- El uso de operadores permite especificar al comando la acción que deseamos realizar.

Arithmetic		Logical		Relational (numeric and string)	
+	addition	&	and	>	greater than
-	subtraction		or	<	less than
*	multiplication	!	not	>=	> or equal
/	division	~	not	<=	< or equal
^	power			==	equal
-	negation			!=	not equal
+	string concatenation			~=	not equal

- De esta manera:

```
browse price rep78 weight length if price<8000
```

```
list make if rep78 == 5 | mpg > 25
```

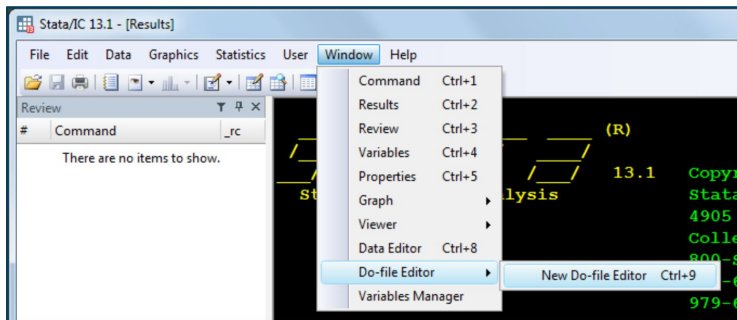
- 1 Introducción
- 2 Interfaz
- 3 Comandos
- 4 Principales comandos I
- 5 Introducción a la programación**
Do-file

- 1 Introducción
- 2 Interfaz
- 3 Comandos
- 4 Principales comandos I
- 5 Introducción a la programación
Do-file**

Do-file

- Para poder manejar datos, hacer cálculos estadísticos y estimaciones de todo tipo, se debe manejar el lenguaje de programación de STATA.
- Sin embargo, cada vez que trabajamos en el entorno principal de STATA y vamos ejecutando comandos, estos se van perdiendo.
- El archivo que nos permitirá programar se denomina **Do-file Editor**, el cual genera un archivo **Do**. Este archivo contiene la programación que manipula base de datos y se pueden ejecutar varias veces.
- Una vez abierto el do file editor, podemos realizar la escritura del programa que podemos ejecutar parcial o totalmente.

Do-file



Principales comandos II

Principales comandos 2	
sort	Ordenar datos de menor a mayor
gsort	Ordena datos de mayor a menor (signo +)
label	Permite darle una etiqueta a cada variable
generate	Describe contenido de datos en memoria o sobre el disco
replace	Reemplaza observaciones de variable
table	Crea tabla
tabulate	Tabulación de variable.
tabstat	Crea tabla con principales estadísticos descriptivos.

Fin