

INTRODUCCIÓN

Evaluaciones: Existen tres evaluaciones centrales al proceso de aprendizaje:

- Evaluación de diagnóstico: Objetivo es ajustar los contenidos y metodología de la clase. 20 estudiantes nunca han ocupado Stata y 4 nunca han ocupado bases de datos cuantitativos y no saben distinguir entre distintos tipos de variables. 20 personas no sabe estimar correlaciones en Stata.
- Tareas: Objetivo es consolidar y profundizar lo aprendido en clases.
- Evaluación final: Objetivo es indicar cuánto les falta aprender para enfrentar el primer trimestre del MPP con éxito.

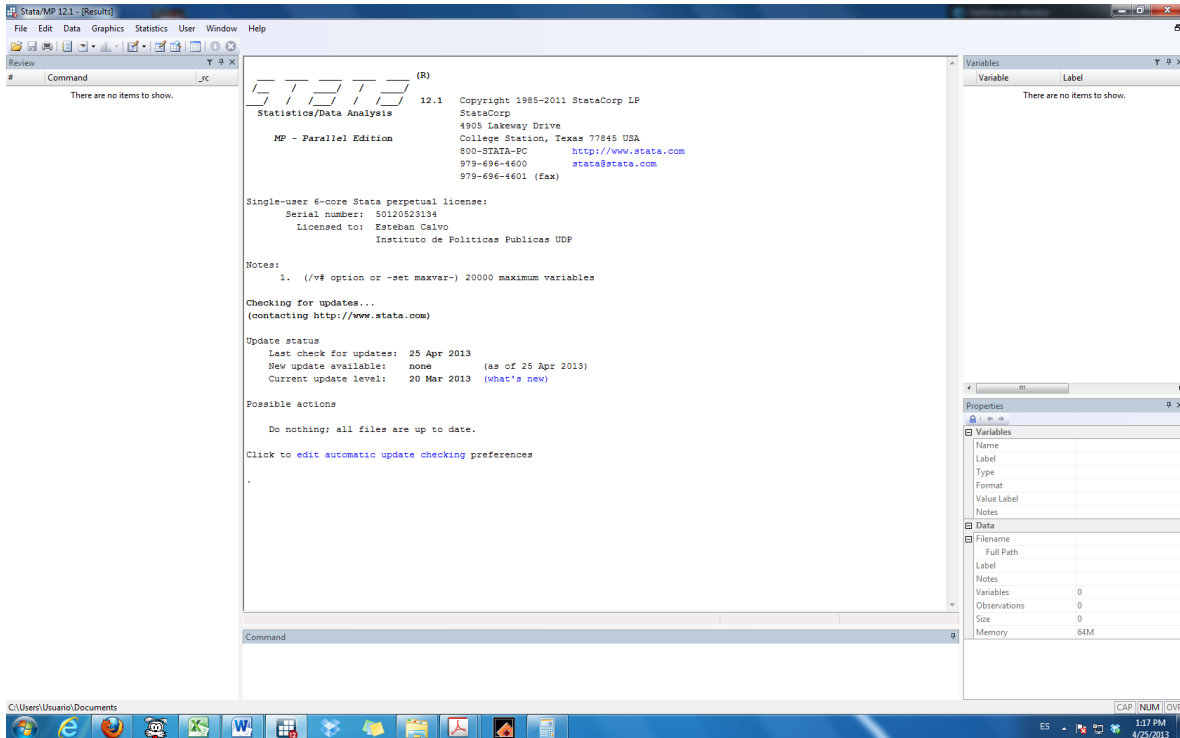
Datos y Variables: Los datos son información (ej: números, texto, símbolos) que puede ser almacenada y analizada. Los datos que no varían pueden ser representados con una constante (ej: el número 5), pero los datos que cambian necesitan ser representados por una variable que puede tomar distintos valores (ej: la variable nota tiene valores que van de 1 a 7). Existen distintos tipos de variables:

Nominal	Dicotómica o binaria	Dos categorías sin un orden claro (0=hombre;1=mujer)
	Categorica	Múltiples categorías sin orden (1=soltero; 2=conviviente; 3=casado; 4=separado; 5=anulado; 6=divorciado; 7=viudo)
Ordinal	Ordinal	Múltiples categorías ordenadas, pero la diferencia entre ellas no es comparable (1=indigente; 2=pobre; 3=no pobre)
Intervalo	Continua	Múltiples valores con una diferencia comparable entre ellos (peso de 50kg; 60kg; 70kg; 80kg; 90kg; 100kg)
	Conteo	Múltiples valores con una diferencia comparable, pero más compleja de lo que parece (0 hijos; 1 hijo; 2 hijos; 3 hijos)

Tipos Básicos de Análisis Estadísticos: Los análisis estadísticos pueden ser clasificados según el número de variables analizadas: univariados (ej: promedio de una variable), bivariados (ej: correlación entre dos variables) o multivariados (ej: regresión múltiple). También pueden ser clasificados según su propósito: describir la población estudiada (ej: número de personas mayores observadas en el censo) o hacer inferencias sobre la población a partir de una muestra (ej: a partir de diferencia observada en el apoyo a un candidato presidencial en una encuesta de opinión a 2.000 personas, generalizar a todos los chilenos).

Stata: Stata es un programa computacional que permite administrar y analizar datos. El nombre viene de la combinación de las palabras statistics y data. Aprender a usarlo puede ser difícil en un principio, pero es una inversión que les abrirá oportunidades y permitirá ahorrar mucho tiempo en el futuro. Dado que seguirán usando Stata a lo largo del MPP, es muy recomendable instalar la última versión de Stata en sus computadores personales. Las licencias para estudiantes pueden ser de 6 meses, anuales o perpetuas y se obtienen a través

de Software Shop. Nuestra ejecutiva de contacto es Janeth Vallejo, janeth@software-shop.com.



Al abrir Stata ves una pantalla con distintos paneles que puedes reubicar a tu gusto:

1. Command (abajo): Permite ingresar los comandos.
2. Results (centro): Registra todos los comandos ejecutados y resultados obtenidos. Las notas aparecen en azul, los errores en rojo y todo el resto en negro. No es posible borrar nada del panel central.
3. Review (izquierda): Registra los comandos ingresados durante una sesión de Stata y permite ir construyendo un código automatizado. Los errores aparecen en rojo y pueden ser borrados.
4. Variables (derecha, arriba): Describe las variables de la base de datos abierta.
5. Properties (derecha, abajo): Describe la base de datos abierta.

Los íconos superiores con celdas o el menú Window permiten abrir el editor de datos en una nueva ventana. Los gráficos también aparecen en una ventana aparte.

Estructura de los Comandos: Stata es un programa basado en comandos, mediante los cuales el usuario le dice lo que quiere hacer. La estructura típica de los comandos está compuesta de cuatro elementos: comando, nombre, condiciones y opciones. Para saber qué decirle a Stata, necesitamos responder cuatro preguntas que corresponden a los cuatro elementos mencionados:

1. Comando: ¿Qué quiero hacer?

2. Nombres: ¿Cuáles son los nombres de las variables o archivos que quiero ocupar?
3. Condiciones: ¿Qué observaciones deseas ocupar?
4. Opciones: ¿Tienes otros requisitos respecto a este programa?

Muchas veces es más fácil especificar con palabras normales lo que quiero hacer y luego intentar pedirselo a Stata. Para los usuarios avanzados, una de las ventajas de Stata es que permite a los usuarios crear sus propios comandos.

Cada comando se debe ingresar en una línea separada, salvo que se esté trabajando en un archivo .do y se utilice /// al final de la línea para indicar que el comando sigue en la línea siguiente.

COMANDOS DE STATA REVISADOS EN ESTA CLASE

Agregar comentarios:

<*comentario> el asterisco permite ingresar un comentario

</*comentario */> señala el comienzo y final de un comentario (solamente en archivo .do)

Preparar Stata:

<update all>actualiza Stata (requiere tener privilegios de administrador del computador)

<set memory, perm>configura la memoria asignada a Stata (no es necesario en versiones recientes del programa; opción perm hace que el cambio sea permanente)

<set maxvar, perm> define el número máximo posible de variables en la base de datos (Stata IC y Small están fijados en 2.048 y 99 variables, respectivamente)

<set matsize, perm>define el número máximo de variables que pueden haber en un modelo (Stata Small está fijo en 100)

<set scrollbufsize> determina cuánto puedes retroceder en la ventana de resultados

<set type *type*, perm> define la precisión con que se almacenan las nuevas variables (*type* puede ir desde byte hasta double, siendo double los números más extensos)

Archivos .log:

Son registros de las operaciones realizadas que se guardan con la extensión .log.

<log using *filename*.log [, append replace]> abre el archivo .log (la opción append agrega el contenido del nuevo log a un log preexistente, mientras que la opción replace reemplaza el log preexistente con el nuevo log)

<log close> cierra el archivo .log y lo guarda automáticamente

<translate> cambia el tipo de archivo entre .log (cualquier editor de texto) y .smcl (Stata)

Abrir y cerrar bases de datos .dta:

<use *filename*.dta, clear> abre un archivo .dta (la opción clear cierra los archivos abiertos)

<save *filename*.dta, replace> guarda un archivo .dta (la opción replace reemplaza archivos preexistentes con el mismo nombre)

Ver y editar los datos:

<browse> permite explorar los datos

<edit> permite editar los datos

Preservar y restaurar los datos:

<preserve> preserva una copia de los datos en la memoria temporal de Stata

<restore [, not]> restaura los datos guardados en la memoria temporal de Stata (la opción not hace que Stata cancele el comando preserve previamente ejecutado)

Cambiarle el nombre a una variable:

<rename *oldvar newvar*> renombrar una variable

Recursos para continuar aprendiendo Stata:

1. Obtener ayuda en Stata (es de poca ayuda para los principiantes):
 - <help *comando*> despliega la ventana de ayuda para un comando específico
 - <search *keyword*> busca palabras claves en la documentación de Stata
 - <net search *keyword*> busca comandos en internet (no pueden instalarse en los computadores de la UDP que requieren privilegios de administrador)
2. Obtener ayuda en el foro Statlist (ojo con preguntar algo vago o ya resuelto):
www.stata.com/statalist/
3. Recursos de UCLA (clases con videos son muy buenas para partir):
<http://www.ats.ucla.edu/stat/stata/>
4. Recursos de UNAV (clases con videos en español):
http://www.unav.edu/departamento/preventiva/recursos_bioestadistica
5. Clases de Natalia Sarkisian (aplicaciones a técnicas estadísticas multivariadas):
<http://www.sarkisian.net/#courses>

DEMOSTRACIÓN PRÁCTICA

```
. * Para copiar los resultados de Stata en Word sin que pierdan su formato, es
    conveniente ocupar letra Courier New Tamaño 10

. * Todo lo que viene después de un asterisco (*) es un comentario que no se ejecuta. Si
    te olvidas del asterisco se produce un error comentario sin asterisco

. * Abrir Stata y ajustar la pantalla según su preferencia

. * Actualizar Stata (esto no funciona en los computadores del laboratorio UDP que
    requieren privilegios de administrador)

. update all
(contacting http://www.stata.com)

Update status
  Last check for updates:  08 Apr 2015
  New update available:    none          (as of 08 Apr 2015)
  Current update level:    23 Jan 2014   (what's new)

Possible actions

  Do nothing; all files are up to date.

. * Descargar la base de datos desde www.esteban-calvo.com

. * Elegir preferencias cuidando de asignar suficientes recursos para que stata pueda
    procesar los datos (algunas preferencias se ajustan automáticamente en versiones
    más nuevas de Stata).

. set memory 800m, perm
set memory ignored.
  Memory no longer needs to be set in modern Statas; memory adjustments are performed
  on the fly automatically.

. set maxvar 20000, perm
(set maxvar preference recorded)

. set matsize 2000, perm
(set matsize preference recorded)

. set scrollbufsize 500000
(set scrollbufsize will take effect the next time you launch Stata)

. set type double, perm
(set type preference recorded)

. *Crear un .log file (prefiero .log para abrirlo en cualquier computador con cualquier
    procesador de texto, idealmente Crimson Editor)

. log using "C:\Users\Esteban\Desktop\clase_1.log", replace
```

```
name: <unnamed>
log: C:\Users\Esteban\Desktop\clase_1.log
log type: text
opened on: 8 Apr 2015, 16:19:44
```

```
. *Abrir la base de datos

. use "C:\Users\Esteban\Dropbox\Public\vs_chile_2005_v9.dta", clear

. *Describir la base de datos
```

```
. desc
```

```
Contains data from C:\Users\Esteban\Dropbox\Public\vs_chile_2005_v9.dta
obs:          1,000
vars:          64          12 Sep 2013 08:10
size:         105,000
```

```
-----
storage display      value
variable name  type  format      label      variable label
-----
yearint       %10.0g
tobscyint     %10.0g
namea         str5   %9s          namea
swb           double %10.0g
lsat          byte   %19.0g      a170
fsat          byte   %19.0g      c006
hap           byte   %16.0g      hap_1
shlth         byte   %10.0g      shlth_1
female        byte   %10.0g      female_1
age           byte   %10.0g
mstat         byte   %10.0g      mstat_1     RECODE of x007 (marital status)
```

```
--Break--
```

```
r(1);
```

```
.. *Utilizar la letra q (o botón break) para pedirle a Stata que no siga desplegando resultados
```

```
. *Utilizar el comando browse para mirar los datos en la ventana del editor
```

```
. browse
```

```
. *Utilizar el comando edit para cambiar los datos en la ventana del editor
```

```
. edit
```

```
. *Guardar una copia de los datos en la memoria temporal de Stata
```

```
. preserve
```

```
. rename namea nombrea
```

```
. edit
```

```
. replace nombrea = "chilequerido" in 1
```

```
nombrea was str5 now str13
(1 real change made)

. replace lsat = -100 in 1
(1 real change made)

. *Recuperar la copia de los datos guardada en la memoria temporal de Stata

. restore

. *Buscar ayuda sobre un comando

. help restore

. *Guardar los datos

. save "C:\Users\Esteban\Desktop\base_clase1.dta", replace
(note: file C:\Users\Esteban\Desktop\base_clase1.dta not found)
file C:\Users\Esteban\Desktop\base_clase1.dta saved

. *Cerrar el log

. log close
name: <unnamed>
log: C:\Users\Esteban\Desktop\clase_1.log
log type: text
closed on: 8 Apr 2015, 16:21:13
```

TAREA PARA LA SIGUIENTE CLASE

Ver los tres primeros videos de IDRA-UCLA (entering, exploring, and modifying data):
<http://www.ats.ucla.edu/stat/stata/notes/default.htm>

Como alternativa se pueden revisar los 10 primeros videos de UNAV (STATA 0.0 a STATA 1.7). Son los mismos tópicos anteriores descritos en español:
http://www.unav.edu/departamento/preventiva/recursos_bioestadistica

Imprimir las notas de las clases antes de ver el video e intentar replicar los procedimientos realizados por la profesora. Guardar el .log file y enviarlo al ayudante al menos una hora antes de la siguiente clase.