

Seminario de STATA

Santiago Garganta¹

Año 2022

El objetivo de este seminario es introducir los conceptos fundamentales del software STATA que se utiliza en varios cursos de la Maestría en Economía de la UNLP. El seminario se desarrolla en cuatro clases durante las cuales se aplican diferentes comandos y herramientas básicas del software, utilizando bases de datos diversas. Las clases se realizarán de forma virtual (vía la plataforma Zoom) y, en caso de ser necesario, se incluirá una clase adicional de consulta para evacuar dudas particulares y generales de los diversos temas abordados en el curso.

La evaluación del curso se realiza mediante un examen práctico domiciliario obligatorio. Por su parte, la asistencia a cada una de las clases será opcional para aquellos que posean los conocimientos básicos de STATA que se dictan en cada una de ellas. A continuación se presenta el programa del curso, el cronograma estimativo y el contenido mínimo de las clases.

Programa y cronograma de clases

Clase 1:

1. Introducción a STATA: interface, sintaxis general, modos de trabajar con STATA, diferentes bases de datos.

2. Comandos básicos: apertura de la base de datos (formato .dta), descripción del contenido de la misma. Creación de variables. Modificación del contenido de una variable, operadores lógicos y sentencias condicionales. Variables dummies y variables aleatorias. Ordenar, renombrar y eliminar variables. Opciones de los comandos.

3. Análisis de datos: cómputo de estadísticas descriptivas, ponderadores, correlaciones, tabulación de datos.

¹ Ante cualquier consulta escribir a: santiago.garganta@gmail.com o sgarganta@cedlas.org

Para prescindir de (la asistencia a) la **Clase 1** es necesario conocer el funcionamiento de los siguientes comandos (con sus diferentes opciones de uso): *clear, cd, use, describe, summarize, return list, sort, gsort, preserve-restore, generate, replace, rename, compare, egen* (funciones: *total, rowtotal, group, tag*), *table (contents), tabulate, tabstat, by, bysort, drop, display, save (saveold)*. Asimismo, se requiere el conocimiento de la sintaxis para la aplicación de sentencias condicionales ([if exp]: ==, !=, >, <), operadores lógicos (&, |), y ponderadores ([weight]).

Clase 2:

4. Desarrollo de un proyecto: programación mediante archivos *do*, archivos *log* e interacción entre los mismos, configuración de directorios del proyecto, y organización de los archivos.

5.1. Gestión de base de datos: importación de datos externos (formatos *.xls, .csv, .txt*) y exportación de una base de datos. Datos experimentales. Etiquetar variables y valores de las variables. Tipo de datos (numérico, *string, elapsed dates*) y modificación del tipo de dato de una variable.

Para prescindir de (la asistencia a) la **Clase 2** es necesario conocer el contenido de la Clase 1, y el funcionamiento de los siguientes comandos adicionales (con sus diferentes opciones de uso): *import delimited (import excel), export delimited (export excel), label variable, label define, label values, destring, encode*, y las funciones *mdy, date, real, string* y *runiform()* aplicadas con el comando *generate*. La organización de un proyecto a través de los archivos *.do* y *.log* también forma parte del contenido mínimo de la Clase 2.

Clase 3:

5.2. Gestión de base de datos: Unión y combinación de bases de datos. Comandos para reestructurar la organización de los datos.

6. Análisis de regresión: estimación por mínimos cuadrados ordinarios, interpretación de los resultados, comandos post-estimación, obtención de valores de coeficientes, y tests de hipótesis. Herramientas gráficas y opciones adicionales útiles en el análisis de regresión.

Para prescindir de (la asistencia a) la **Clase 3** es necesario conocer el contenido de las clases 1 y 2, y el funcionamiento de los siguientes comandos adicionales (con sus diferentes opciones de uso): *merge, append, xpose, reshape (wide/long)*,

collapse, regress, predict (opciones: *xb* y *residuals*), operadores "*_b[varname]*" y "*_se[varname]*", *test, estimates, outreg2, graph twoway* (*scatter/line*).

Clase 4:

7. Otras herramientas y comandos: macros locales, variables y archivos temporales, diferencias entre los comandos *include* y *do*, ejecución de archivos *.do* con argumentos, alternativas para almacenar resultados, matrices.

Para prescindir de (la asistencia a) la **Clase 4** es necesario conocer el contenido de las clases 1, 2 y 3, y el funcionamiento de los siguientes comandos adicionales (con sus diferentes opciones de uso): *local, matrix, smat, mkmat, do, include, args*.

Bibliografía

Baum, C. (2006). *An introduction to modern Econometrics using Stata*. Stata Press (capítulos 1, 2 y 4).

Hamilton, L. (2013). *Statistics with Stata (Updated for Version 12)*. Brooks/Cole-CENGAGE Learning.

Stata (2007). *User's Guide*. Stata Press.

Stata (2013). *STATA Base Reference Manual. Release 13*. Stata Press.