

2020-1

1. Horario atención a estudiantes

Profesor: Cristhian Julian Acosta Pardo
Horario de clase: miércoles de 5:00 p.m. a 6:20 p.m.
Salón: ML - 207
Horario de atención: Jueves 4:00 p.m a 5:00 pm, con cita previa vía correo electrónico.
Lugar de atención: W - 826

2. Introducción y descripción general del curso

"It is not only the violin that shapes the violinist, we are all shaped by the tools we train ourselves to use, and in this respect programming languages have a devious influence: they shape our thinking habits." - Edsger W. Dijkstra

Stata es un software ampliamente utilizado en la investigación económica y en ciencias sociales. En este curso se busca que los estudiantes adquieran comprensión general de las herramientas y su uso común para la investigación. A lo largo del curso, los estudiantes practicarán situaciones que surgen frecuentemente en el manejo de datos y su análisis cuantitativo, estas deben resolverse de manera trazable, verificable y replicable. El énfasis será sobre la construcción y procesamiento de bases de datos, la presentación de estadísticas descriptivas y resultados econométricos por medio del uso de diferentes herramientas.

Se busca que los estudiantes entiendan la estructura general de programación en *Stata* como complemento a los conocimientos adquiridos en otras clases que ya hacen uso de este software. Se esperan dos resultados al final del curso, que los estudiantes se sientan cómodos creando rutinas automatizadas y eficientes para tareas recurrentes en la investigación en economía y que sean capaces de encontrar soluciones a problemas complejos de programación en *Stata* de manera autónoma.

3. Objetivos

Proporcionar a los estudiantes herramientas para el manejo y uso de bases de datos en *Stata* con enfoque de investigación en economía.

- A) Entender la estructura de información utilizada en la investigación en economía a través de uso de los comandos de *Stata*
- B) Introducir y aplicar los principios de programación en *Stata* para automatizar tareas comunes de manera trazable, verificable y replicable.
- C) Desarrollar soluciones a problemas complejos de programación de manera autónoma para procesar bases de datos y comunicar resultados de análisis estadísticos.

4. Organización del curso

Módulo 1: Manipulación de Bases de Datos

Enero 22 – Clase 1: Introducción, importación y descripción de datos.

- Tipos de archivos
- Sintaxis de un comando
- Importar datos de otros formatos a Stata
- Comandos para comprimir, organizar y guardar bases de datos (e.g. *compress*, *order*, *sort*)

Enero 29 – Clase 2: Variables I

- Generación de variables (e.g. *gen*)
- Nombres y listas de variables (e.g. *rename*, *replace*)
- Operadores lógicos (i.e. condicionales)
- Comandos para la descripción de variables (e.g. *sum*, *tab*, *tabstat*, *describe*)
- Variables numéricas
- Variables de texto
- Conversión entre variables (e.g. *tostring*, *destring*)

Febrero 5 – Clase 3: Variables II

- Etiquetas
- Rutinas *by*, *bysort*
- Generación de variables (e.g. *egen*)
- Missing Values
- Recodificación de variables (e.g. *recode*)
- Funciones para variables de texto (*string*)

Febrero 12 – Clase 4: Variables III - Fechas y tiempo

- Tipos de fechas
- Conversión de fechas
- Manipulación de fechas

Febrero 19 – Clase 5: Unión y compresión de bases de datos

- *append*
- *merge*
- *collapse*
- *contract*
- *reshape*

Febrero 26 – Práctica en clase - Taller 1

➤ Fecha límite de entrega del Taller 1: viernes 13 de marzo

Módulo 2: Automatización de tareas repetitivas

Marzo 4 – Clase 6: Escalares, vectores y matrices

- *scalar*
- *matrix define*
- Manipulación de escalares, vectores y matrices.
- Matrices como instrumentos para guardar información: *mkmat*, *svmat*

Marzo 11 – Clase 7: Macros

- Local
- Global
- if

➤ Semana de receso: marzo 16 – marzo 20

➤ Fecha límite para la entrega del 30%: viernes 20 de marzo

➤ Último día para solicitar retiros: marzo 27, 6:00 p.m.

Marzo 25 – Clase 8: Loops

- *While*
- *foreach*
- *forvalues*

Abril 1 – Clase 9: Loops II

- Anidación
- Comandos para programar: *tokenize*, *levelsof*

Abril 15 – Práctica en clase - Taller 2

➤ Fecha límite de entrega del Taller 2: viernes 1 de mayo

Módulo 3: Presentación de resultados

Abril 22 – Clase 10: Gráficos I - Clasificación

- Introducción
- Gráficos básicos
- Gráficos *twoway*

Abril 29 – Clase 11: Gráficas II

- Edición de gráficos *twoway*
- Unión de Gráficos

Mayo 6 – Clase 12: Gráficos III: Mapas

- Coordenadas geográficas
- Lectura y proyección de mapas
- Mapas coropléticos

Mayo 13 – Clase 13: Exportación de Resultados Estadísticos

- *outreg2*
- *outsheet*
- *tabout*
- *putexcel*

Mayo 20 – Práctica en clase - Taller 3

➤ Fecha límite de entrega del Taller 3: jueves 4 de junio

5. Metodología

El curso se realizará en salas de computadores y tiene un importante componente práctico. Durante la clase, el profesor explicará los conceptos del tema y los ilustrará por medio de ejemplos desarrollados en Stata, en su mayoría con datos públicos. El curso cuenta con notas de clase que se entregarán al inicio de cada sesión, estas son el soporte de lo explicado respecto al análisis de datos en *Stata*.

En algunas sesiones los estudiantes desarrollarán ejercicios como complemento a la explicación. En otras, los estudiantes desarrollarán ejercicios entregables al final de la clase con supervisión. El estudio autónomo de los conceptos se realizará a través del taller de cada módulo del curso. Cada taller se inicia en clases dirigidas y se entrega terminado dos semanas después. Todas las preguntas adicionales sobre el material de la clase, material complementario, ejercicios en clase y los enunciados de los talleres se responderán en los horarios de atención.

Además de las salas de computadores de la universidad, para acceder a la última versión del software los estudiantes pueden utilizar la plataforma *Nukak* (<https://nukak.uniandes.edu.co/>), ingresando con su usuario y contraseña a la “Sala Virtual”.

Fechas importantes:

Viernes 20 de marzo	Fecha de entrega del 30% de la nota del curso
Viernes 27 de marzo	Último día para retiro de cursos
Sábado 23 de mayo	Último día de clases del semestre

6. Competencias

Al final del curso, los estudiantes podrán:

- Manejar do-files y programación en Stata.
- Analizar datos para obtener información relevante.
- Presentar resultados estadísticos de forma adecuada.

7. Evaluación

La evaluación del curso se realizará mediante los ejercicios en clase y los talleres. El do-file con la respuesta de cada actividad es lo único que se debe enviar a través de Sicua en la fecha establecida en el programa. La calificación final será determinada de la siguiente manera:

- Ejercicios en clase: 25%
- Talleres: 25% c/u

Cada pregunta en los talleres se calificará con la siguiente rúbrica de calificación:

Puntos	Descripción
Todos	El código corre independientemente, está bien y sigue las instrucciones del taller
Mitad	El código corre independientemente, está bien, pero no se siguieron las instrucciones (ej: no usar loops, no utilizar el comando automatizado, etc.)
Mitad	El código no corre independientemente porque se cometió un error en un paso anterior pero individualmente está bien.
Ninguno	El código no corre independientemente, no se da respuesta a la pregunta o la respuesta está mal desarrollada.

Las notas de los talleres estarán disponibles 10 días hábiles después de la entrega.

Reclamos: Según los artículos 62 y 63 del Reglamento general de estudiantes de pregrado, el estudiante tendrá cuatro días hábiles después de la entrega de la evaluación calificada para presentar un reclamo. El profesor responderá al reclamo en los cinco días hábiles siguientes. Si el estudiante considera que la respuesta no concuerda con los criterios de evaluación podrá solicitar un segundo calificador al Consejo de la Facultad en los cuatro días hábiles posteriores a la recepción de la decisión del profesor.

Fraude: Se considera fraude cuando, en una actividad, dos estudiantes presentan el mismo código, la secuencia de comandos es idéntica y no es posible identificar el proceso de aprendizaje de cada uno. En caso de detectar un fraude se seguirán los lineamientos establecidos en el Reglamento general de estudiantes de pregrado.

8. Bibliografía y material de apoyo

Además de las notas de clase, se recomienda consultar los siguientes recursos:

- Cameron, A. & Trivedi, P. (2009). Microeconometrics Using Stata. Stata Press.
- Essam, T. & Hughes, A. (2016). Fundamentals of data analysis and visualization [En: <http://geocenter.github.io/StataTraining/>]
- Herrera, M. (2015). Econometría espacial usando Stata. Breve guía aplicada para datos de corte transversal. Universidad Nacional de Salta, Argentina.
- Lachine, I. et al. Stata Reference Manual. What you should know about Stata after taking the Stata introduction course.
- Mitchell, M. (2004). A Visual Guide to Stata Graphics. Stata Press.
- Pisati, M. (2014) The A to Z of how to create thematic maps of Italy using spmap. Italian Stata Users Group meeting.
- Software Collections at IDEAS: <https://ideas.repec.org/i/c.html>
- Stack Overflow (Preguntas y respuestas sobre programación): <https://stackoverflow.com/>
- Stata documentation: <https://www.stata.com/features/documentation/>
- Stata resources for learning : <https://www.stata.com/links/resources-for-learning-stata/>
 - CheatSheets : <https://www.stata.com/bookstore/stata-cheat-sheets/>
 - Stata Tutorial: <https://data.princeton.edu/stata>
- Stata FAQ: <http://www.stata.com/support/faqs/>
- Statalist (Preguntas y respuestas sobre programación en Stata): <https://www.statalist.org/>
- The Stata Journal: <https://www.stata-journal.com/>
- UCLA guide to Stata: <http://www.ats.ucla.edu/stat/stata/>