

1. Información del equipo pedagógico y horarios de atención a estudiantes

Profesor magistral: Cristhian Julián Acosta Pardo

Correo electrónico: cj.acosta@uniandes.edu.co

Atención a estudiantes: viernes de 10:00 a.m. a 11:00 a.m. Vía [Zoom](#)

2. Descripción 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 habilidades para su uso común en el análisis de datos, en el ejercicio profesional y la investigación. A lo largo del curso los estudiantes experimentarán situaciones que surgen frecuentemente en el manejo de datos y su análisis cuantitativo que deben abordarse de manera trazable, verificable y replicable. El énfasis será sobre la construcción y procesamiento de bases de datos y la presentación de resultados de análisis sencillos.

Se busca que los estudiantes entiendan la estructura general de la programación para el análisis de datos, complemento a lo realizado en otras clases, donde ya hacen uso de algún software estadístico, como *Stata*. Se esperan que al final del curso los estudiantes se sientan cómodos creando rutinas automatizadas y eficientes para el análisis de datos frecuentes en economía y que sean capaces de solucionar problemas complejos de automatización de rutinas de manera autónoma. Con estas habilidades, se espera que los estudiantes sean capaces de aprender otros lenguajes de programación en el futuro, tomando como base los conceptos aprendidos en este taller, según las necesidades de su desarrollo profesional.

3. Resultados de aprendizaje

- Escribe secuencias de órdenes en el lenguaje de *Stata-Mata* para limpiar datos y hacerlos aptos para generar análisis exploratorios de datos, procesando adecuadamente valores atípicos y considerando valores faltantes.
- Desarrolla rutinas de operaciones con estructuras de programación por casos o iteradas que permiten automatizar procesamiento de datos y operaciones comunes del análisis exploratorio de datos.
- Genera gráficos o tablas de resultados de procesos de análisis exploratorio de datos que permiten describir un comportamiento de interés para contribuir a la respuesta de una pregunta de negocio o la indagación de hipótesis.

4. Cronograma

Módulo 1: Manipulación de Bases de Datos

Clase 1: Introducción, importación y descripción de datos.

- Interfaz, Tipos de archivos, Sintaxis de un comando
- Cargar, explorar, organizar y guardar bases de datos.
- Importar bases de datos de otros formatos a Stata.

Clase 2: Variables I

- Comandos para la descripción y exploración de variables
- Generación, nombres y listas de variables
- Operadores lógicos (i.e. condicionales)
- Variables numéricas y de caracteres
- Conversión entre variables

Clase 3: Variables II

- Etiquetas
- Tratamiento de valores faltantes (*Missing Values*)
- Recodificación de variables
- Funciones para variables de caracteres
- Generación de variables con funciones adicionales

Clase 4: Variables III - Fechas y tiempo

- Tipos de fechas
- Conversión y manipulación de fechas

Clase 5: Unión y compresión de bases de datos

- Unión de bases: vertical y horizontal
- Agregación de bases de datos
- Cambio de estructura - *Reshape*

Práctica en clase – Taller 1: 28 de febrero - Fecha de entrega: 17 de marzo

Módulo 2: Automatización de tareas repetitivas

Clase 6: Escalares, vectores y matrices

- Manipulación de escalares, vectores y matrices.
- Matrices como instrumentos para guardar información: *mkmat*, *svmat*
- Escritura básica en MATA

Clase 7: Macros

- Local, Global
- Condicionales
- Funciones extendidas
- Archivos temporales

Semana de receso y semana santa: marzo 18 – 31

Clase 8: Iteraciones I

- Con condicional
- Sobre listas y números

Clase 9: Iteraciones II

- Indexación de elementos (tokenize)
- Anidación

Práctica en clase - Taller 2: 17 de abril – Fecha de entrega: 13 de mayo

Módulo 3: Presentación de resultados

Clase 10: Gráficos I - Clasificación

- Introducción
- Gráficos básicos: Barras
- Gráficos twoway: Dispersión, tendencia, histogramas.

Clase 11: Gráficas II

- Edición y unión de gráficos twoway
- Creación de gráficos con iteraciones
- Gráficos especiales con programación: Barras con intervalos de confianza

Clase 12: Gráficos III: Mapas

- Coordenadas geográficas
- Lectura y generación de mapas
- Mapas coropléticos: categorización de variables
- Mapas con objetos

Clase 13: Exportación de Resultados Estadísticos

- Exportación de estadísticas descriptivas
- Exportación de resultados de regresiones
- Creación de tablas de resultados personalizadas

Práctica en clase - Taller 3 - Fecha de entrega: 4 de junio

5. Referencias

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/>]
- Gabor, B. & Gabor, K. (2021). Data Analysis for Business, Economics, and Policy. Cambridge University Press. New York. [<https://gabors-data-analysis.com>]
- 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](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/](https://www.statalist.org/)
- The Stata Journal: <https://www.stata-journal.com/>
- UCLA guide to Stata: <http://www.ats.ucla.edu/stat/stata/>

6. Metodología

El curso se realizará con base en el aprendizaje basado en problemas y la indagación autónoma. Para esto, cada semana tendrá dos momentos:

1. El estudio de videos elaborados por el profesor en los que se desarrollan los aspectos técnicos de los temas de la semana. En su conjunto, los videos tienen una duración máxima de una hora y veinte minutos. Este material está disponible en el repositorio en [GitHub](#) o en Bloque Neón, y debe ser estudiado en su totalidad antes de la clase por los estudiantes. El profesor explica los temas de la semana a través de ejemplos con herramientas en *Stata*, en su mayoría con datos públicos y basados en problemas reales. También se suministran los *do-files*, los datos y las notas de clase, que son el soporte intuitivo sobre el análisis de datos en *Stata*. Las notas de clase contienen explicaciones prácticas y conceptuales para el manejo de datos con *Stata* y no reemplazan la documentación del software, que debe ser consultada constantemente.
2. En la clase presencial se abordarán los aspectos conceptuales y prácticos del tema de la semana. Se responderán las preguntas de los estudiantes sobre el material del repositorio y se desarrollará un ejercicio, con el acompañamiento del profesor, para ser entregado al finalizar la semana. En cualquier momento de la clase el profesor puede hacer un quiz, a modo de control de estudio del material del repositorio, que hará parte de la calificación del ejercicio de la semana. Las clases no serán grabadas.

Esta metodología tiene en cuenta las siguientes consideraciones:

El aprendizaje de un lenguaje de programación requiere de práctica autónoma. Con los videos se busca que el estudiante pueda replicar la ejecución de las rutinas, entendiendo su funcionamiento y la intuición detrás de ellas, con la oportunidad de repetir temas en los que tuvieron dificultad y moderar el ritmo de su estudio. Estos videos se convierten en material de consulta directa, adaptando el desarrollo del curso a las condiciones particulares de los estudiantes.

- Se requiere de un alto nivel de compromiso y disciplina para seguir los temas del curso cada semana. De esta forma las clases presenciales son más provechosas, compartiendo preguntas, opiniones, generando discusiones y desarrollando los ejercicios en clase.
- Los temas se dividen en tres módulos. En la última semana de cada módulo se realizará un procesamiento de datos en un Taller. Previo a la clase de la semana se entregará un enunciado que debe ser estudiado. En la sesión sincrónica se responden preguntas sobre el enunciado, se discuten sobre posibles soluciones propuestas por los estudiantes.

La Universidad cuenta con licencias campus de *Stata* a las que se puede acceder a través del siguiente [enlace](#) con su correo electrónico institucional y siguiendo las instrucciones de descarga e instalación del software en su computador personal (para [Windows](#) y [Mac](#)). También se puede acceder a la última versión del software a través de la plataforma [Nukak](#), ingresando con su usuario y contraseña a la “Sala Virtual”.

7. Evaluaciones

La evaluación del curso se realizará mediante la revisión de un ejercicio en clase por semana y tres talleres durante el *semestre*. El do-file con la respuesta de cada actividad es lo único que se debe enviar a través de *Bloque Neón* en la fecha y hora establecida en el programa. No se tendrán en cuenta entregas tardías sin justificación o excusa. La calificación final será determinada de la siguiente manera:

Ejercicios en clase + quiz: 25%

Talleres: 25% cada uno

La calificación puede estar acompañada de un comentario personal o al grupo en general que apunta a corregir errores o complementar lo entregado. Cada pregunta en los talleres se calificará con la siguiente guía de *calificación*:

Puntos	Descripción
Completo	El código ejecuta independientemente, está bien escrito y soluciona el problema y cumple con los objetivos.
Mitad	El código se ejecuta independientemente pero no se siguieron las instrucciones, no se cumple el objetivo o la solución del problema no es completa.
Mitad	El código no se ejecuta independientemente porque se cometió un error en un paso anterior que no permite llegar a la solución del problema.
Ninguno	El código no corre independientemente por problemas de sintaxis, no se da solución al problema. No hay desarrollo.

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

Justificación de ausencia: Según el artículo 45 del Reglamento General de Estudiantes de Pregrado, los estudiantes tendrán ocho días hábiles para presentar una excusa válida y, de ser aceptada, el profesor programará el supletorio en las dos semanas siguientes. Tenga en cuenta que en el art. 45 se especifican un mínimo de excusas válidas, pero también se especifica que “el profesor podrá tener en cuenta otras circunstancias que a su criterio puedan justificar la ausencia del estudiante”.

Reclamos: Según los artículos 64, 65 y 66 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 o más 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 artículo 116 del Reglamento General de Estudiantes de Pregrado.

8. Asistencia

Según los artículos 43 y 44 del Reglamento General de Estudiantes de Pregrado (RGEPr), los estudiantes pueden no asistir hasta al 20% de las sesiones. Si la inasistencia supera dicho umbral el profesor seleccionará la nota de ejercicio de clase más alta para ser reemplazada por 0.

9. Políticas generales de los cursos de Economía y fechas importantes

Los estudiantes deben consultar [este enlace](#), donde se encuentran las reglas sobre asistencia a clase, excusas válidas, fraude académico y faltas disciplinarias, reclamos, políticas de bienestar y fechas importantes del semestre.