

1 Equipo Pedagógico y Horario de Atención

Profesor: David Santiago Caraballo Candela.

Correo electrónico: ds.caraballo@uniandes.edu.co

Horario de Clase: Jueves de 5:00 p.m. a 6:20 p.m.

Salón de Clase: Edificio ML-107.

Horario de Atención: Viernes de 2:00 p.m. a 3:00 p.m. Vía [Zoom](#) (con cita previa).

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 estará en 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 espera 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** 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 procesamientos 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: Exploración y procesamiento 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: Limpieza de datos.

- Criterios de limpieza.
- Tipos de datos, etiquetas, recodificación de variables.
- Unidad de observación.
- Valores faltantes.

Clase 3: Exploración de datos I.

- Generación de variables con funciones **egen**.
- Frecuencias, histogramas.
- Estadísticas descriptivas.

Clase 4: Exploración de datos II.

- Uso de datos con fechas.
- Manejo de valores atípicos.

Clase 5: Agregación por agrupación de datos.

- Cálculos agregados.
- Cálculos agrupados.
- **preserve-restore**.

Clase 6: Uniones verticales y horizontales.

- Casos de recolección y unión de datos: ¿cuándo usar la unión vertical u horizontal?
- Añadir observaciones: unión vertical.
- Añadir variables y cruzar tablas: unión horizontal.

Clase 7: Limpieza de tablas con **reshape**.

- Formas anchas y largas (repaso).
- Transformaciones de tablas y usos.

Taller 1: Enunciado 18 de septiembre. [Entrega 28 de septiembre.](#)

Módulo 2: Automatización de tareas repetitivas

Clase 8: Análisis por casos: Macros y condicionales.

- Macros locales y globales.
- Condicionales.
- Estudio autónomo: Funciones extendidas y archivos temporales.

Clase 9: Procesamiento con ciclos iterados I.

- Iteraciones sobre listas de archivos.
- Práctica: Uniones verticales.

Clase 10: Procesamiento con ciclos iterados II.

- Práctica: Uniones horizontales.
- Iteraciones sobre listas de variables.
- Iteraciones anidadas.

Clase 11: Exploración con ciclos iterados.

- Revisión de variables.
- Generación de tablas.
- Generación de gráficos.

Taller 2: Enunciado 16 de octubre. Entrega 2 de noviembre.

Módulo 3: Presentación de resultados

Clase 12: Análisis con variables categóricas y gráficos **oneway**.

- Introducción a los gráficos.
- Gráficos de barras.
- Gráficos para estadísticas descriptivas.

Clase 13: Análisis de relación entre variables continuas y gráficos **twoway**.

- Gráficos de dispersión, tendencia, histogramas.
- Sintaxis para la creación de gráficos compuestos.
- * Tema opcional: Gráficos especiales con programación.
- * Tema opcional: Mapas con **Stata**.

Clase 14: Edición de gráficos para presentación.

- Títulos y textos.
- Colores, formas.
- Etiquetas de ejes.
- Creación de gráficos con iteraciones.

Clase 15: Exportación de resultados estadísticos.

- Exportación de gráficos.
- Exportación de tablas con estadísticas descriptivas.

Clase 16: Creación de reportes personalizados a Excel.

- Diseño del reporte.
- Programación del reporte.
- Automatización del reporte.

Taller 3: Enunciado 13 de noviembre. Entrega 30 de noviembre.

5 Metodología

El curso plantea en cada clase una serie de situaciones del quehacer con datos en el ejercicio profesional del economista. El hilo conductor está determinado por el proceso de análisis exploratorio de datos, para establecer la limpieza, el procesamiento, el análisis y la generación de reportes a partir de una necesidad ya definida.

El curso tiene dos (2) espacios:

1. La clase presencial práctica, donde se abordarán los aspectos conceptuales y prácticos del tema de la semana. En todas las clases se desarrolla un código que debe completarse como tarea y enviarse a través de Bloque Neón, como ejercicio de la semana.
2. Material en línea para el estudio autónomo de temas según la necesidad de análisis. Se dispone de un repositorio en [GitHub](#) con videos elaborados por el profesor en los que se desarrollan los aspectos técnicos del uso de los comandos y herramientas más utilizadas en **Stata**. Estos corresponden a una versión anterior a este curso y son un material de consulta. Los videos se organizan por “clases”, y cada clase tiene una duración máxima de una hora y veinte minutos. Para cada “clase” 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.

Esta metodología considera que el aprendizaje de un lenguaje de programación y del análisis de datos 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.

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 computadora 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”.

6 Criterios de Evaluación

La evaluación del curso se realizará mediante la revisión de un (1) ejercicio en clase por semana y tres (3) talleres en pareja durante el semestre. **El do-file con la respuesta de cada actividad (incluyendo las preguntas del .pdf) es lo único que se debe enviar a través de Bloque Neón en la fecha y hora establecida.** No se tendrán en cuenta entregas tardías sin justificación o excusa válida. La calificación final será determinada de la siguiente manera:

Talleres en pareja (3)	25% cada uno
Ejercicios en clase	25% promedio

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 rúbrica de calificación:

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

6.1 Asistencia y Excusas

Los estudiantes podrán no asistir hasta al 20% de las sesiones del curso. El profesor tiene autonomía para definir la consecuencia de inasistencia sin excusa válida. **En particular, esta constará de la reprobación automática del curso con nota de uno coma cinco (1,5).**

Según el Artículo 45 del [Reglamento General de Estudiantes de Pregrado](#), los estudiantes tendrán tres (3) días hábiles para presentar una excusa válida y, de ser aceptada, el profesor programará el supletorio en las dos semanas siguientes.¹ Tenga en cuenta que en el Artículo 45 también 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.” **Todas las excusas provistas por los estudiantes pasarán por un proceso de verificación interno por parte de la Facultad de Economía.**

6.2 Reclamos

Según los Artículos 64, 65 y 66 del [Reglamento General de Estudiantes de Pregrado](#), el estudiante tendrá cuatro (4) días hábiles tras conocer las calificaciones en cuestión para presentar un reclamo de forma escrita. El reclamo debe ser enviado por correo electrónico al profesor que incluya un documento ordenado en el que anexe imágenes de la evaluación (si así fuere conveniente) y una descripción detallada del reclamo. El profesor responderá al reclamo en los cinco (5) 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 de Economía en los cuatro (4) días hábiles posteriores a la recepción de la decisión del profesor.

6.3 Fraude

El fraude, en cualquiera de sus formas, no es admisible bajo ninguna circunstancia. Cualquier evidencia de fraude será remitida al Comité Disciplinario del Consejo de la Facultad de Economía, a través del cual los estudiantes involucrados deberán proceder a remitir sus descargos. Concluido el proceso disciplinario, la evaluación o actividad académica respectiva podrá ser calificada, a discreción del profesor, hasta con nota de cero coma cero (0,0), entendida como la consecuencia académica y sin perjuicio de la sanción disciplinaria impuesta.

En este sentido, 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 manera individual. Detalles adicionales disponibles en el Artículo 116 del [Reglamento General de Estudiantes de Pregrado](#).

7 Cláusulas

7.1 Ajustes razonables y momentos difíciles

En el siguiente [enlace](#) se encuentra la información sobre las políticas de ajustes razonables y momentos difíciles admisibles por la Universidad. Estas políticas buscan apoyar a los estudiantes en situaciones excepcionales. Para solicitar alguno de los dos, el estudiante debe contactar al profesor de forma oportuna y, de ser necesario, informar a la Decanatura de Estudiantes.

¹Se consideran excusas válidas: a) incapacidades médicas; b) incapacidades expedidas por la Decanatura de Estudiantes; c) muerte del cónyuge o de parientes hasta el segundo grado de consanguinidad o de afinidad; d) autorización para participar en eventos deportivos, expedida por la Decanatura de Estudiantes; e) autorización para asistir a actividades académicas y culturales, expedida por la respectiva dependencia académica; y f) citación a diligencias judiciales, debidamente respaldada por el documento respectivo (véanse la reglamentación de las incapacidades estudiantiles y el Acuerdo 126 del Consejo Académico, sobre participación estudiantil en eventos académicos y deportivos).

7.2 Respeto por la diversidad

Todos debemos respetar los derechos de quienes integran esta comunidad académica. Consideramos inaceptable cualquier situación de acoso, acoso sexual, discriminación, matoneo o amenaza. Cualquier persona que se sienta víctima de estas conductas puede denunciar su ocurrencia y buscar orientación o apoyo ante alguna de las siguientes instancias: el equipo pedagógico del curso, la Coordinación o la Dirección del programa, la Decanatura de Estudiantes, la Ombudsperson o el Comité MAAD. Si requiere más información sobre el protocolo MAAD establecido para estos casos, puede acudir a Nancy García Duarte (n.garcia@uniandes.edu.co) en la Facultad de Economía. Más información sobre el protocolo MAAD disponible en el siguiente [enlace](#).

7.3 Herramientas de IA

El uso de la inteligencia artificial (IA) en las actividades académicas debe ser informado, transparente, responsable, ético, crítico y coherente con las instrucciones de las actividades académicas de cada curso, con los reglamentos de la Universidad y con los valores de libertad, integridad, excelencia y solidaridad de la Universidad de los Andes. Todos los miembros de la comunidad deben consultar el siguiente [enlace](#) para interiorizar sus buenas prácticas.

8 Referencias

- Cameron, A. & Trivedi, P. (2009). *Microeconometrics Using Stata*. Stata Press.
- ★ ChatGPT. Recuperado de: <https://chatgpt.com/g/g-ouo55RZmd-stata-analyst>
- ★ CheatSheets. Recuperado de: <https://www.stata.com/bookstore/stata-cheat-sheets/>
- Essam, T. & Hughes, A. (2016). *Fundamentals of data analysis and visualization*. Recuperado de: <http://geocenter.github.io/StataTraining/>
- Gabor, B. & Gabor, K. (2021). *Data Analysis for Business, Economics, and Policy*. Cambridge University Press, New York. Recuperado de: <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.
- Mitchell, M. (2004). *A Visual Guide to Stata Graphics*. Stata Press.
- Software Collections at IDEAS. Recuperado de: <https://ideas.repec.org/i/c.html>
- Stack Overflow. Recuperado de: <https://stackoverflow.com/>
- Stata Documentation. Recuperado de: <https://www.stata.com/features/documentation/>
- Stata Tutorial. Recuperado de: <https://data.princeton.edu/stata>
- Stata FAQ. Recuperado de: <http://www.stata.com/support/faqs/>
- Statalist. Recuperado de: <https://www.statalist.org/>
- The Stata Journal. Recuperado de: <https://www.stata-journal.com/>
- UCLA guide to Stata. Recuperado de: <http://www.ats.ucla.edu/stat/stata/>

Fecha de entrega del 30%: 10 de octubre

Último día para retirar cursos de 16 semanas: 12 de septiembre

Último día para subir notas finales a Banner: 12 de diciembre