

## Análisis Estadístico con Stata.

### Curso presencial

#### Objetivo del curso:

Al finalizar el curso, el participante tendrá la capacidad de:

Manipular una base de datos para obtener estadísticos descriptivos. Fusionar, colapsar y recodificar bases de datos como paso previo para la estimación de modelos. Emplear Stata 14 tanto en el modo interactivo y a través de archivos de ejecución por lotes (do-files) para obtener estadísticos descriptivos, gráficos, estimaciones y post estimación de modelos. Elegir entre varias técnicas de estimación según el tipo de problema teórico que enfrente.

#### Requisito(s):

Tener conocimiento básico de estadística.

#### Descripción del curso:

El curso “Análisis estadístico con Stata” tendrá una duración de 24 horas, distribuidas en 4 sesiones de 6 horas cada una. Es un curso práctico en el cual se manejarán diversas bases de datos. Las primeras sesiones serán dedicadas al manejo del programa y al empleo de comandos para la preparación de las bases de datos previo al análisis descriptivo y la estimación. Posteriormente se presentarán las metodologías de estimación y post- estimación con su respectiva aplicación a un conjunto de datos.



Imparte: Miguel Ángel Cruz  
Especialista en Economía Aplicada

Miguel Ángel Cruz es especialista en Econometría Aplicada por la Universidad Nacional Autónoma de México y cuenta con una licenciatura en Economía por la Universidad Nacional Autónoma de México. En su actual cargo como Especialista en Software Estadístico para MultiON Consulting, desarrolla webinars y cursos presenciales de programas de análisis estadístico.

## Temario

1. **Introducción e inicio de sesión en Stata.**
  - 1.1 Inicio de sesión
  - 1.2 Estructura de Stata
    - 1.2.1 Results
    - 1.2.2 Graph
    - 1.2.3 Viewer
    - 1.2.4 Command
    - 1.2.5 Review
    - 1.2.6 Variables
    - 1.2.7 Properties
    - 1.2.8 Data Editor
    - 1.2.9 Do-file Editor
    - 1.2.10 Variables Manager
  - 1.3 La barra de herramientas de Stata
  - 1.4 Configuración del tamaño de la memoria
  - 1.5 Configuración del número de variables
  - 1.6 Configuración del número máximo de variables en un modelo
  - 1.7 Directorios de trabajo
  - 1.8 Manuales y documentación
    - 1.8.1 El sistema de ayuda
  - 1.9 Variables y observaciones
    - 1.9.1 Variables
    - 1.9.2 Observaciones
  - 1.10 Entrada de datos
    - 1.10.1 Cargar archivos de datos a memoria
    - 1.10.2 Modo directo para capturar datos
    - 1.10.3 Importar archivos de datos
    - 1.10.4 Cargar archivos con formato de Stata
  - 1.11 Exportar datos
    - 1.11.1 Exportar a XML
    - 1.11.2 Exportar a ASCII
  - 1.12 Valor de etiqueta (Value labels)
  - 1.13 Formato de despliegue
2. **Comandos de Stata.**
  - 2.1 Comandos básicos
  - 2.2 Sintaxis general
    - 2.2.1 Save
    - 2.2.2 Append
    - 2.2.3 Unión de datos

- 2.2.3.1 Unión de datos uno a uno
  - 2.2.4 Compress
  - 2.2.5 Copy
  - 2.2.6 By
  - 2.2.7 If ,&, |, In
  - 2.2.8 Clonevar
  - 2.2.9 Codebook
  - 2.2.10 Edit
  - 2.2.11 Erase
  - 2.2.12 Memory
  - 2.2.13 Inspect
  - 2.2.14 Update
  - 2.3 Interface con el sistema operativo
3. Manejo de archivos y datos.
- 3.1 Manipulación de datos
    - 3.1.1 Generate
    - 3.1.2 Operadores relacionales y lógicos
    - 3.1.3 Replace
    - 3.1.4 Encode
    - 3.1.5 Decode
    - 3.1.6 Destring
    - 3.1.7 Tostring
    - 3.1.8 Egen
    - 3.1.9 Order
    - 3.1.10 Sort
    - 3.1.11 Set obs
  - 3.2 Desplegado de los datos
  - 3.3 Resumen de datos
  - 3.4 Cambiar y renombrar variables
    - 3.4.1 3.4.1 Drop
  - 3.5 Archivos Log
4. Comandos básicos de estadística.
- 4.1 Count
  - 4.2 Summarize
  - 4.3 Means
  - 4.4 Centile
  - 4.5 Pctile
  - 4.6 Correlate
  - 4.7 Regress
  - 4.8 Tablas

- 4.8.1 Table
- 4.8.2 Tabulate
  - 4.8.2.1 Tabulate one way
  - 4.8.2.2 Tabulate two way
- 4.8.3 Drawnorm

## 5. Introducción a la Programación.

- 5.1 Archivo do
- 5.2 Program define
- 5.3 Program define en un archivo do
- 5.4 Combinación de archivos do

## 6. Manejo de Gráficos y Matrices.

- 6.1 Gráficos
  - 6.1.1 Gráficos de dispersión
  - 6.1.2 Histogramas
  - 6.1.3 Gráficos de caja
  - 6.1.4 Matrices de gráficas de dispersión
  - 6.1.5 Gráficas de funciones definidas
- 6.2 Manejo de matrices
  - 6.2.1 Input matriz by hand
  - 6.2.2 Matriz input y matrix
  - 6.2.3 Define matrix from expresión
  - 6.2.4 Convert variables to matrix
  - 6.2.5 Convert matrix to variables
  - 6.2.6 Nombrar los renglones y columnas de una matriz
  - 6.2.7 Extraer elementos de una matriz
  - 6.2.8 Operadores de matrices
    - 6.2.8.1 Transpuesta de una matriz
    - 6.2.8.2 Insertar una matriz debajo de otra
    - 6.2.8.3 Insertar una matriz a la derecha de otra
    - 6.2.8.4 Suma de matrices
    - 6.2.8.5 Resta de matrices
    - 6.2.8.6 Multiplicación de matrices
    - 6.2.8.7 División de una matriz entre un escalar
  - 6.2.9 Funciones de matrices
    - 6.2.9.1 Identidad
    - 6.2.9.2 Inversa
    - 6.2.9.3 Traza
  - 6.2.10 Vectorización de una matriz
  - 6.2.11 Determinante de una matriz

## 7. Análisis de Regresión Lineal.

- 7.1 Análisis de residuales
- 7.2 Prueba de normalidad
  - 7.2.1 Prueba de Shapiro-Wilk
  - 7.2.2 Causas que dan origen a la falta de normalidad
  - 7.2.3 Gráfica de normalidad
- 7.3 Gráfica de residuales vs valores ajustados
- 7.4 Gráfica de residuales vs la variable de predicción

## 8. Análisis de Regresión Lineal Múltiple.

- 8.1 Estimación del modelo
- 8.2 Significancia del modelo y de los parámetros
- 8.3 Análisis de residuales
  - 8.3.1 Prueba de normalidad
  - 8.3.2 Gráfica de normalidad
  - 8.3.3 Gráfica de residuales vs valores ajustados
- 8.4 Outliers y puntos de influencia (PI) en regresión lineal
  - 8.4.1 Outliers
  - 8.4.2 Leverage
  - 8.4.3 Puntos de influencia
    - 8.4.3.1 Distancia de Cook
    - 8.4.3.2 DFFITS
    - 8.4.3.3 DFBETAS

## 9. Análisis de Series de Tiempo.

- 9.1 Series de tiempo en Stata
- 9.2 Gráficas de series de tiempo
- 9.3 Componentes de una serie de tiempo y su correlograma
  - 9.3.1 Tendencia
  - 9.3.2 Estacionalidad
  - 9.3.3 Estacionariedad
- 9.4 Operadores en series de tiempo
  - 9.4.1 Rezago  $L$
  - 9.4.2 Diferencia Estacional  $S$
  - 9.4.3 Diferencia  $D$
- 9.5 Ajuste de un modelo de series de tiempo
  - 9.5.1 Procesos autorregresivos
  - 9.5.2 Procesos de media móvil
- 9.6 Predicción en series de tiempo

## 10. Análisis Multivariado

- 10.1 Análisis de componentes principales (ACP)

- 10.2 Análisis de conglomerados (clusters)
  
- 11. Regresión Logística
  - 11.1 Introducción
  - 11.2 Conceptos fundamentales
    - 11.2.1 Regresión logística
    - 11.2.2 El modelo de probabilidad lineal y sus problemas
    - 11.2.3 Modelos logit y probit
  - 11.3 Modelo logit
  - 11.4 Modelo probit
  - 11.5 Diagnóstico del modelo
    - 11.5.1 Estadísticas de clasificación, sensibilidad y especificidad del modelo estimado
    - 11.5.2 Curva ROC
    - 11.5.3 Gráfica de sensibilidad y especificidad
  
- 12. Ejercicios y Casos de los Participantes
  - 12.1 Ejemplo adicionales
    - 12.1.1 Ejemplo 1
    - 12.1.2 Ejemplo 2
    - 12.1.3 Ejemplo 3
  - 12.2 Ejercicio



Cómputo Científico y Técnico: *software y hardware especializado.*



MultiON es líder en México y Latinoamérica en la comercialización, el soporte y la capacitación en el uso de software y hardware especializado para la ciencia, la educación, la industria y los servicios. Esta empresa de cómputo científico y técnico fue fundada en 1989 por el ingeniero y maestro en administración Joaquín Antonio Maury González durante sus estudios de doctorado y es 100% mexicana.

## Joel Cervantes

Asesor Comercial LATAM

MultiON Consulting, S.A. de C.V.

Cómputo Científico y Técnico: *software y hardware especializado.*

Tel: +52 (55) 5559-4050 Ext. 119 | [cursos@multion.com](mailto:cursos@multion.com)

[www.multion.com](http://www.multion.com)

