



Investigación Social Aplicada Usando STATA

DESCRIPCIÓN

Objetivo: Brindar a los alumnos las herramientas necesarias para el manejo, análisis descriptivo y exploratorio de bases de datos en ciencias sociales, sean de micro datos o de datos agregados. En particular, se busca que los alumnos del curso desarrollen la capacidad para desenvolverse en el entorno de Stata 12 para (i) administrar y consolidar las bases de datos que tengan a su disposición, resolviendo de manera independiente los problemas usualmente presentes en bases de datos de ciencias sociales (e.g. tratamiento de missing values, uso de factores de expansión, corrección de errores sistemáticos en los datos, construcción de variables de interés); (ii) describir y analizar exploratoriamente los datos, en búsqueda de patrones presentes en ellos mediante el uso de las herramientas del software (e.g. distribuciones, gráficos, mapas geográficos, tablas, tests); y (iii) realizar algunas pruebas de hipótesis introductorias, según sean necesarias y frecuentes en cada ciencia social¹. A lo largo del curso, se hará énfasis en que los alumnos “aprendan a aprender” en Stata 12, la que suele ser la habilidad más importante en el software, mediante el uso adecuado del archivo de ayuda, el manual y los servidores especializados de usuarios del software.

Público objetivo: Investigadores o tesisistas en ciencias sociales y alumnos de últimos ciclos de facultad con interés en la investigación con algún tema de investigación en mente, y que requieren el manejo de información cuantitativa, poco familiarizados con Stata.

Estrategia metodológica:

- Modalidad presencial para las presentaciones en clases prácticas.

- Manejo interactivo de los casos desarrollados en el curso, tomando en cuenta los principales temas de interés del conjunto de estudiantes.
- Presentación, y análisis exploratorio y descriptivo de las principales bases de datos nacionales y globales. Se hará énfasis en aquella información de acceso fácil y público, cuya información pueda ser estudiada, en principio, desde diferentes ciencias sociales; no obstante, se atenderán consultas sobre bases de datos específicas de interés para los alumnos.
- Evaluación en base a un trabajo (para casa) aplicando las herramientas desarrolladas en el curso. La aprobación de la evaluación es necesaria para obtener el certificado del curso.

Conocimientos previos recomendados: Estadística (Probabilidad y Estadística Inferencial).

Certificación: Se entregará certificado de participación a quienes hayan asistido al 90% de las sesiones.

Informes: cisepa@pucp.edu.pe

EXPOSITOR

Cesar Huaroto. Estudiante de Doctorado en Economía en la Universidad Nacional de La Plata. Magister y Licenciado en Economía por la PUCP. Con amplia experiencia en temas de Econometría de Evaluación de Impacto aplicada a Desarrollo Económico. Particularmente interesado en los tópicos relacionados a Pobreza, Desigualdad, Educación y Recursos Naturales.

Con investigaciones presentadas en diversas conferencias locales e internacionales y en revistas académicas arbitradas y tengo experiencia de docencia tanto en cursos de pregrado y posgrado en PUCP y UARM. Ha trabajado en consultorías para Fondoempleo, el Ministerio de Desarrollo (MIDIS) y para Apoyo Consultoría. Manejo avanzado de paquetes estadísticos como STATA y ArcGIS.

Campos de especialización: Microeconometría, Evaluación de Impacto, Pobreza.

PROGRAMA

Sesión 1	Elementos básicos de Stata y del manejo de bases de datos.	<ul style="list-style-type: none"> ○(1a) Presentación. ¿Por qué Stata? ¿Qué pueden decirnos los datos? ○(1b) “A preñdiendo a hablar en Stata”. Algunas convenciones de Stata. La interface de Stata y sus elementos. El uso de do-files y log-files para ordenar nuestro trabajo. ○(1c) Manejo de bases de datos en ciencias sociales. Acceso a e ingreso de los datos. Descripción y exploración de la base de datos. Creación y modificación de variables. Tipos de variables: string y numéricas. Missing values. Manejo de etiquetas. Formas long y wide de las bases de datos. Juntando bases de datos. Almacenamiento y exportación de la base de datos.
Sesión 2	Análisis descriptivo I.	<ul style="list-style-type: none"> ○(2a) Distribuciones: Tablas. Valores medios de una distribución: media, mediana, y moda. Dispersión y sus medidas: desviación estándar, varianza, coeficiente de variación. Tabulaciones y tablas. ○(2b) Distribuciones: Gráficos. Gráficos para distintos tipos de variables: categóricas, categóricas ordenadas, cuantitativas. Manejo y almacenamiento de gráficos. Otros gráficos en Stata. Series de tiempo.
Sesión 3	Análisis descriptivo II.	<ul style="list-style-type: none"> ○(3a) Estadísticos descriptivos. Uso de tablas en Stata. Uso apropiado de factores de expansión: tipos de factores de expansión. Diseños de encuesta. Pruebas de hipótesis: ¿Qué nos dicen y cómo las interpretamos? ¿Cuándo podemos confiar en ellas? ¿Cómo las ejecutamos? ○(3b) Mapas en Stata. Construcción de mapas en Stata: acceso a shapefiles, transformación y uso de comandos pertinentes. Diseño de mapas: distritales, provinciales, departamentales, por cuenca, internacionales. Manejo y almacenamiento de mapas.
Sesión 4	Análisis exploratorio.	<ul style="list-style-type: none"> ○(4a) Relaciones entre dos o más variables 1. Gráficos de dispersión. Correlaciones, covarianzas. Correlación entre

		<p>variables ordinales. Regresión bivariada y multivariada: supuestos del modelo clásico de regresión lineal, “usos y abusos”, alcances y limitaciones.</p> <p>o(4b) Relaciones entre dos o más variables 2. Análisis de varianza: one-way y two-way. Análisis factorial: principal factor y principal-component factor.</p>
Sesión 5	Elementos básicos de programación y solución de problemas prácticos.	<p>o(5a) Programación funcional en Stata. Loops: -foreach- y -forvalues-. Macros: -local- y -global-. Manejo de los contenidos de r(.) y e(.). Matrices. Comandos -egen-. Prefijo -by-.</p> <p>o(5b) El problema de los missing values y las soluciones al problema. La naturaleza del problema. Imputación: supuestos, métodos y comandos relevantes.</p>
Sesión 6	Aplicaciones prácticas	<p>o(6a) Repaso de todo lo revisado durante clase.</p> <p>o (6b) El helpfile y otros recursos para el aprendizaje. El archivo de ayuda de Stata: alcances y limitaciones. Statalist. Comandos en -ssc install-. Ejercicios.</p> <p>o(6c) Presentación. Investigación aplicada usando Stata, usando todo lo aprendido en una investigación.</p>

Referencias bibliográficas

- Acock, A.C.
2010 “A gentle introduction to Stata”, 3rd Ed. College Station, TE : Stata Press, 2010.
- Baum, C.
2005 “A little bit of Stata programming goes a long way...”. SUGUK 2005, invited lecture. Disponible en <http://fmwww.bc.edu/ec-p/wp612.pdf>.
- 2009 “An introduction to Stata programming”. College Station, TE : Stata Press,
2009 2011 “Using Stata for Applied Research: Reviewing its Capabilities” En: Journal of Economic Surveys, 2011, 25:2, 380-394
- StataCorp
2011 “Stata 12 Base Reference Manual”. College Station, TX: Stata Press.