

**-CURSO DE EDUCACIÓN PERMANENTE 2020-**

## **Recursos, usos y estrategias para el tratamiento sistemático de datos en proyectos y trabajos de estudiantes**

**Equipo docente: Fernando Filgueira (docente responsable), Fabricio Carneiro, Nicolás Schmidt, Jimena Pandolfi, Santiago López Cariboni y Gustavo Méndez**

- Desde el 1 de setiembre hasta el 13 de octubre de 2020.
- Martes y viernes de 16 a 18 horas
- Modalidad: 10 sesiones virtuales y 3 talleres presenciales complementarios (15, 22 y 29 de setiembre)
- A través de la plataforma Zoom.
- Carga horaria: 20 horas.
- Matrícula: sin costo.
- Dirigido a egresadas/os universitarios, estudiantes de posgrado y estudiantes de grado avanzadas/os..
- [Formulario de inscripción](#).

### **Objetivos:**

Este curso busca familiarizar al estudiante con los recursos disponibles en la Unidad de Métodos y Acceso a Datos y ejemplificar diferentes estrategias para explotar dichos recursos mediante módulos que se presentarán a lo largo del curso. Asimismo para abordar dichos ejemplos se ofrece una breve introducción a temas de diseño metodológico y herramientas de análisis. Al finalizar el curso el estudiante tendrá un conocimiento exhaustivo de los recursos disponibles en la UMAD, la estructura de un subconjunto clave de bases de datos y la introducción a herramientas para su tratamiento.

### **Contenidos:**

#### **Sesión 1: Un paseo virtual por los recursos básicos de la UMAD (Fernando Filgueira y equipo UMAD)**

- Observatorio del Uruguay
- Sistema de búsqueda personalizado
- Acceso a micro-datos y bases
- Linkoteca de recursos en materia de sistemas de información, bases y búsquedas personalizadas
- Los recursos para la producción de datos: centro de llamadas, sala gessel y laboratorio de experimentos
- Recursos pedagógicos, cursos online y descripción del vector metodológico de la Facultad de Ciencias Sociales

**Consigna:** Elegir cualquiera de los recursos disponibles para redactar y presentar datos sobre uno de los siguientes temas en Uruguay: evolución de la pobreza monetaria y de las NBI, los resultados de las elecciones 2019 en Uruguay y su comparación con las elecciones anteriores en primera y segunda vuelta, la economía Uruguaya en el último quinquenio (2015-2019): crecimiento, inversión y competitividad; Dinámicas de población recientes en el Uruguay: fecundidad, natalidad y arreglos familiares.

## **Sesión 2: Las tradiciones epistémicas en la investigación social y el uso de los números en las mismas**

- La tradición analítico-deductiva: consistencia formal teórica y derivación a tests empíricos
- La tradición cuantitativa tradicional: el modelo experimental y sus emulaciones no experimentales. La pureza del modelo experimental puro. La imperfección de los modelos cross-sectional y la complejidad de los modelos no experimentales de series temporales agrupadas.
- La tradición cuantitativa configuracional: Causalidad compleja, fuzzy sets, modelos descriptivos de conglomerados y factoriales.
- La tradición histórico comparada y los métodos mixtos. Los datos cuantitativos como soportes empíricos estilizados y análisis causal no inspirado en la tradición experimental
- El experimento de laboratorio: las ciencias del comportamiento en settings experimentales puros.

**Consigna de trabajo:** breve ejercicio de múltiple opción e intercambio sobre las respuestas correctas y las proporcionadas por los estudiantes

## **Sesión 3: Estructura básica de la información y elementos claves de la estadística.**

- La estructura tripartita de la información. Una mirada desde el diseño.
- Que es una variable, una unidad y un valor. Exhaustividad y exclusión.
- Muestra y población.
- El teorema central del límite y la ley de los grandes números.
- Las curvas de distribución y parámetros de las mismas: medias, moda mediana, varianza y desvío estándar.
- El concepto de independencia estadística y la noción de correlación. Que quiere decir que una correlación es estadísticamente significativa.
- Tres formas de observar correlación: comparación de medias, chi cuadrado, regresión lineal.

**Consigna:** Se entrega una matriz de estructura básica de la información con tres variables. Construir desde esa información una curva de distribución de frecuencias simples de la primer variable, establecer su media, moda, mediana, varianza y desvío estándar. Construir una segunda curva de frecuencias o distribución con la segunda variable y hacer lo mismo. Generar un gráfico de tipo scatter y la regresión lineal que surge de combinar estas dos variables. Comparar las medias de ambas variables a partir de la tercer variable tricotómica ordinal. Interpretar.

#### **Sesión 4: Microdatos, metadatos, formulario y diccionario de variables. Un ejemplo desde las encuestas de hogares en Uruguay (Fernando Filgueira y Jimena Pandolfi)**

- Encuestas de Hogares, Censos, Encuestas de Salud, Encuestas de Opinión Pública. Los formularios como fuentes potenciales de preguntas.
- Generando reportes básicos para conocer nuestras bases.
- La construcción de variables.

**Consigna:** se le entregará una ECH2018 y una serie de operaciones simples preestablecidas a realizar con un set de variables. Se les pedirá describir la información de las salidas.

#### **Sesión 5: Respondiendo preguntas sobre desigualdad, pobreza, empleo, ingresos y vivienda desde las encuestas de hogares. Ejemplos de trabajo con la ECH en Uruguay. (Jimena Pandolfi)**

¿Son los hogares del quintil 1 en promedio más pobres hoy que en 2010? Dos respuestas posibles: en capacidad de compra absoluta, en capacidad de compra relativa a los hogares del quintil 2. Las variables de empleo y su construcción desde la ECH. . Construyendo el ingreso total de los hogares y el ingreso per cápita de los hogares. ¿Que tipo de distribución tiene los ingresos per cápita de los hogares?. ¿En qué quintil y decil de ingresos están sus hogares?

**Consigna:** Generar curva de kernel de distribución de los ingresos y definir porcentaje debajo de la línea de pobreza a partir de datos proporcionados desde la UMAD. Usando la ECH 2015 y 2018 ver tasas de actividad, empleo, desempleo en base a script en STATA y base proporcionada, Comentar los resultados comparando ambos años.

#### **Sesión 6: Construyendo sistemas de información, monitoreo y evaluación de las políticas públicas: el observatorio de los derechos humanos, el monitor de la seguridad social del CINVE, el monitor educativo de primaria y el observatorio educativo del INEED (Fernando Filgueira y Gustavo Mendez).**

Usando registros administrativos y encuestas periódicas para construir sistemas de monitoreo y evaluación de áreas de interés y de políticas públicas. Tres ejemplos: el monitor educativo de primaria, el observatorio de la seguridad social y el observatorio de derechos humanos. Definición de variables que representen aspectos sustantivos de nuestros temas y problemas, fuentes de información, construcción de indicadores.

**Consigna:** Ingresar al observatorio del MIDES, al Monitor educativo de primaria y al Mirador educativo del INEED. Seleccionar tablas y gráficos que permitan discutir y evaluar el problema de la repetición en primer año de escuela entre el años 2000 y el año 2015. Evolución, distribución y posibles causas.

#### **Sesión 7: Análisis electoral. Encuestas de intención de voto, votos, partidos, representación. Series históricas y su análisis. (Nicolas Schmidt y Fabricio Carneiro). Las elecciones departamentales: encuestas de intención de voto.**

¿Quién ganaría las elecciones departamentales en Montevideo, Canelones y Maldonado de acuerdo a las últimas encuestas de opinión?. Encuestas de opinión sobre intención de voto. De las preguntas del formulario al gráfico del informativo. Las matriz original,

los resultados en bruto y los resultados ajustados y sin ajustar. Cómo asignamos casos ambiguos. Qué es el margen de error.

**Consigna:** A definir.

**Sesión 8: Utilizando los sistemas de consulta personalizada para pensar preguntas de investigación y de allí a los micro-datos. Un ejemplo desde los recursos de la UMAD (Fabricio Carneiro).**

¿Es cierto que creemos menos en la democracia? ¿Desde cuando? ¿Hace cuanto?. ¿Todos igual? ¿Empeoramos más que la región?. ¿Eso está asociado al ciclo electoral?. ¿Qué preguntas nos aproximan a medir cuánto creemos en la democracia? Nuestras encuestas de OP, el latinobarómetro, LAPOP.

**Consigna:** haciendo uso del latinobarómetro compare la adhesión a la democracia de Uruguay con el resto de América Latina entre los años 2010, 2014 y 2018.

**Sesión 9: COVID19 y las cuatro formas de validez en los diseños metodológicos: validez externa, validez de constructo, validez estadística y validez interna. (Fernando Filgueira)**

Cuatro tipos de validez que pueden ser también fuentes de ideas para investigaciones. Los diseños muestrales y la validez externa: porque es difícil sostener la equi-probabilidad de selección de unidades en las investigaciones de campo (el ejemplo de COVID y de las encuestas de opinión). La validez de constructo: el problema olvidado? El problema de la validez externa en casos en los cuales no es posible realizar selección aleatoria. La validez interna: el modelo molar/experimental, la validez interna como combinación de mecanismos causales densos y varianza. La validez estadística y el problema de la violación de los supuestos en las técnicas de análisis y el debate sobre que es significativo. Los diseños muestrales y la validez externa: porque es difícil sostener la equi-probabilidad de selección de unidades en las investigaciones de campo (el ejemplo de COVID y de las encuestas de opinión). El problema de la validez externa en casos en los cuales no es posible realizar selección aleatoria. El problema de la validez de constructo para medir el contagio y asignar los decesos.

**Consigna:** Identificar el problema de validez externa y de constructo del indicador de prevalencia a partir de los casos detectados mediante PCR. Cargar la base COVID19 de la UMAD en R y aplicar los scripts para graficar casos acumulados, casos diarios, casos sobre tests diarios y google mobility en Uruguay.

**Sesión 10: Series de tiempo, análisis de series temporales agrupadas. Validez estadística y validez interna. (Santiago Lopez Cariboni)**

¿La menor movilidad (timing, sostenimiento en el tiempo y profundidad) ayuda a explicar la evolución de la pandemia en los países en América Latina? Desagregando los componente de las series temporales y estableciendo un análisis causal. Validez interna y validez estadística.

**Consigna:** Intercambiar con el docente los hallazgos que el presente. Desagregar la serie temporal de casos nuevos por día de Chile.

### **Método didáctico:**

En cada sesión el estudiante deberá leer o asistir/mirar materiales obligatorios (textos, PPTs y videos) acompañados de una presentación del docente y realizar tareas finales de sesión bajo consigna docente. La primera hora (aproximadamente) de cada sesión es expositiva y la segunda es de acompañamiento para el inicio de trabajo en la consigna (s).

### **Sistema de evaluación:**

La nota final del curso se compone de dos elementos: participación en las sesiones (20%), y la compleción satisfactoria de las consignas de trabajo (80%) que en cada sesión el/la docente le solicitará al estudiante. Un conjunto de bases de datos, series, tablas y programas de análisis serán entregados a los estudiantes al inicio del curso y a lo largo de los módulos para la realización de las consignas.

### **Bibliografía:**

#### Bibliografía básica

Cook, Thomas y Campbell, Donald; Quasi-Experimentation. Design & Analysis Issues for Field Settings. Houghton Mifflin Company; Boston/Dallas/London, 1979.

Johan Galtung: "La estructura básica de la información".

Francis Korn, "El significado del término variable en sociología"

Lieberson, S. 1985. Making it count: the improvement of social research and theory. Berkeley: University of California Press.

Wallace, Walter: "La lógica de la ciencia en la sociología", Alianza Universidad, Madrid, 1979. (1, 2, 3 y 9).

Materiales en youtube: La ley de los grandes números y el teorema central del límite, probabilidad, las medidas de asociación.

Aplicación del teorema central del límite a una ejemplo muestral: <https://www.youtube.com/watch?v=o2afi9BKRIM>

Procesos estocásticos para pensar series temporales <https://www.youtube.com/watch?v=pxNVCCgFx1o>

Teorema central del límite y prueba de hipótesis

Materiales con links: <https://umad.cienciassociales.edu.uy/>

Otra bibliografía complementaria

Beck, N. 2008. Time-series cross-sectional analysis. Pp. 475–93 in *The Oxford handbook of political methodology*, edited by J. Box-Steffensmeier, H. Brady, and D. Collier. Oxford: Oxford University Press.

Freedman, D. A. 2010. Black ravens, white shoes, and case selection. In *Statistical models and causal inference: a dialogue with the social sciences*, by David Freedman and edited by D. Collier, J. Sekhon, and P. Stark. Cambridge: Cambridge University Press.

Mill, J. S. 1974/1843. Of the four methods of experimental inquiry. In book 3 chapter 8, *A system of logic*. Toronto: University of Toronto Press.

Ragin, C. 1987. *The comparative method: moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.