

## Programa

Curso	:	<b>Recursos, usos y estrategias para el tratamiento sistemático de datos en proyectos y trabajos de estudiantes</b>
Carácter	:	Semi presencial (teóricos vía zoom, laboratorios presenciales)
Docentes	:	Fernando Filgueira, Jimena Pandolfi, Nicolás Schmidt, Fabricio Carneiro, Gustavo Méndez, Santiago López
Mail	:	umad@cienciassociales.edu.uy
Duración	:	10 sesiones de 2 horas cada una
Lugar	:	Sala de Informática, FCS (Constituyente 1502, 3er piso)
Web	:	<a href="https://github.com/UMAD-FCS/Curso-UMAD">https://github.com/UMAD-FCS/Curso-UMAD</a>

## Objetivos

Este curso busca familiarizar al estudiante con los recursos disponibles en la Unidad de Métodos y Acceso a datos (UMAD) 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 (teórica y práctica) a herramientas para su tratamiento.

## Contenido

### **Sesión 1: Un paseo virtual por los recursos básicos de la UMAD**

Docente: *Fernando Filgueira*

- 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
- Usando registros administrativos y encuestas periódicas para construir sistemas de monitoreo y evaluación de áreas de interés y de políticas públicas. Cuatro ejemplos: Observatorio de la Educación de ANEP, Observatorio Social del MIDES, Observatorio Territorio Uruguay de OPP y el Observatorio de Derechos Humanos. Definición de variables que representen aspectos sustantivos de nuestros temas y problemas, fuentes de información, construcción de indicadores.

#### **Consignas:**

- Ingresar al Observatorio Social del MIDES, al Observatorio de la Educación de ANEP 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 los años 2005 y 2015. Evolución, distribución y posibles causas.
- Se trabajará con los recursos disponibles en la UMAD para compilar y presentar datos sobre algunos de los siguientes temas en Uruguay a elegir: (1) evolución de la pobreza monetaria y de las NBI; (2) los resultados de las elecciones nacionales 1985-2019 y análisis de tendencias; (3) la economía Uruguaya en el último quinquenio (2015-2019): crecimiento, inversión y competitividad; (4) Dinámicas de población recientes en el Uruguay: fecundidad, natalidad y arreglos familiares. El ejercicio implica ingresar al centro de datos del Observatorio Uruguay de la UMAD, seleccionar algún gráfico y/o tabla e interpretarlo.
- Ingresar al Observatorio Territorio Uruguay de OPP. Buscar el indicador de alumnos por maestro en educación primaria. Explorar el metadato y visualizar cómo está construido el indicador. Reconstruir el indicador para el año 2018 a partir de la fuente administrativa: Anuario Estadístico MEC 2018.

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

Docente: *Fernando Filgueira*

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:**

- Breve ejercicio de múltiple opción e intercambio sobre las respuestas correctas y las proporcionadas por los estudiantes.

**Sesión 3 y 4: Estructura básica de la información y elementos claves de la estadística.**

Docente: *Fernando Filgueira*

- La estructura tripartita de la información. Una mirada desde el diseño.
- Qué 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 5: Microdatos, metadatos, formulario y diccionario de variables. Un ejemplo desde las encuestas de hogares en Uruguay.**

Docente: *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 la 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 6: Respondiendo preguntas sobre desigualdad, pobreza, empleo, ingresos y vivienda desde las encuestas de hogares. Ejemplos de trabajo con la ECH en Uruguay.**

Docente: *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 que 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 R y base proporcionada. Comentar los resultados comparando ambos años.

**Sesión 7: Análisis electoral. Encuestas de intención de voto, votos, partidos, representación. Series históricas y su análisis.**

Docente: *Nicolás Schmidt y Fabricio Carneiro*

¿Quien 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. Como asignamos casos ambiguos. Que 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 el los recursos de la UMAD**

Docente: *Nicolás Schmidt y 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?. Que preguntas nos aproximan a medir cuanto 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.**

Docente: *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.**

Docente: *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.

## Metodología

En cada sesión el estudiante deberá leer o asistir/mirar materiales obligatorios (textos, presentaciones 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).

## 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 los largo de los módulos para la realización de las consignas.

## Contactos Administrativos

*Inscripción y consultas:*

Educación Permanente de FCS:  
`educacion.permanente@cienciassociales.edu.uy`

*Certificación y consultas sobre créditos:*

Bedelía de posgrado de FCS:  
`docedu.permanente@cienciassociales.edu.uy`

## Referencias

- BECK, N. (2008). Time-series-cross-section methods. *The Oxford handbook of political methodology* , 475–493.
- COOK, T. D., CAMPBELL, D. T. Y DAY, A. (1979). *Quasi-experimentation: Design & analysis issues for field settings*, vol. 351. Houghton Mifflin Boston.

- FREEDMAN, D. A. (2010). *Statistical models and causal inference: a dialogue with the social sciences*. Cambridge University Press.
- GALTUNG, J. (1973). Teoría y métodos de la investigación social. Tech. rep.
- KORN, F. (1969). El significado del término variable en sociología. *Artículo incluido en Korn F y otros, "Conceptos y .*
- LIEBERSON, S. (1987). *Making it count: The improvement of social research and theory*. Univ of California Press.
- MILL, J. S. (1843). Of the four methods of experimental inquiry. *A System of Logic, Raciocinative, and Inductive* , 450–479.
- RAGIN, C. C. (2014). *The comparative method: Moving beyond qualitative and quantitative strategies*. Univ of California Press.
- TEAM, R. C. (2000). *Introducción a R. Notas sobre R: Un entorno de programación para Análisis de Datos y Gráficos*. CRAN. URL <https://cran.r-project.org/doc/contrib/R-intro-1.1.0-espanol.1.pdf>.
- WALLACE, W. L. (1976). La lógica de la ciencia en la sociología. Tech. rep.
- WICKHAM, H. Y GROLEMUND, G. (2017). *R for Data Science*. O'Reilly Media. URL <https://r4ds.had.co.nz/>.