

Seminario “Técnicas de análisis multivariado en R”

Docente: Dra. María Celeste Ratto

Dra. en Ciencia Política, Investigadora Independiente CONICET en el Instituto de Investigaciones en Diversidad Cultural y Procesos de Cambio (IIDyPCA)
Universidad Nacional de Río Negro/CONICET

Fecha del curso: 23 al 26 de septiembre 2024

Descripción del curso:

El curso estará dedicado al estudio de algunas de las técnicas estadísticas multivariadas más utilizadas en las ciencias sociales. El énfasis se pondrá en los aspectos aplicados de cada técnica en lugar de sus fórmulas y algoritmos computacionales. Se desarrollarán ejercicios con datos reales utilizando el software de libre acceso R. También se trabajará con artículos que muestren ejemplos de cómo emplear las técnicas estadísticas en publicaciones.

Objetivos:

La idea central del curso será tener una visión general sobre el análisis de datos aplicado a la investigación en ciencias sociales. El curso tiene dos objetivos principales. En primer lugar, tiene la intención de proporcionar a los estudiantes los conocimientos y habilidades necesarias para analizar empíricamente datos reales de sus propias investigaciones. Para ello, se brindarán algunas herramientas básicas tanto de análisis de datos, como su aplicación y manejo a partir del programa R. En segundo lugar, quiere darles una visión general de las principales técnicas estadísticas multivariadas utilizadas en las ciencias sociales para comprender y poder discutir los trabajos empíricos de otros autores en su disciplina.

En tal sentido, si bien se explicarán nociones básicas de matemática y estadística que permitan comprender las técnicas empleadas, el seminario se centrará en la cocina de los datos, enfocando principalmente en la práctica y experiencia.

Metodología:

El curso tendrá una duración de 15 horas reloj, en el horario de 18 a 21 por la plataforma Zoom. Cada sesión estará dividida en un primer módulo, durante el cual se explicarán conceptualmente las nociones fundamentales de las técnicas multivariadas y a partir de ejemplos empleados en artículos. Luego en un segundo módulo se desarrollará el laboratorio aplicando lo aprendido a una base de datos, en el programa R. Las y los alumnos contarán con el script de R explicado para poder ejecutar el análisis correspondiente. Cada día se desarrollará una técnica multivariada.

Evaluación:

La evaluación final del curso constará de un artículo corto (máximo 10 páginas) en el cual los estudiantes deberán aplicar una de las técnicas de análisis vistas durante

el curso a alguna base de datos, realizando el análisis mediante R. El artículo deberá contener la presentación del problema, una breve descripción del marco teórico, una descripción de la técnica a emplear, los resultados y su interpretación. Deberán entregar el trabajo, el script de R empleado y la base de datos.

Requisitos:

Encontrarán una guía indispensable para este curso en el libro: *AnalizaR Datos Políticos*, Francisco Urdinez y Andrés Cruz Labrín (editores). 2019; versión 0.1

Podrán acceder al libro mediante este link:

<https://arcruz0.github.io/libroadp/index.html>

IMPORTANTE: Vamos a necesitar que Todxs traigan ya instalados el programa R y R Studio. Eso nos va a ahorrar mucho tiempo al comienzo.

Instalar R es fácil, independiente de si usas Windows, Mac o Linux. Basta con ingresar a <https://cran.r-project.org/> y seguir las instrucciones de descarga e instalación.

Para instalar RStudio, es necesario ya haber instalado R. La descarga e instalación es accesible en Windows, Mac y Linux.

El link es <https://www.rstudio.com/products/rstudio/download/#download>

En el capítulo 2 “R Básico”, del libro *AnalizaR Datos Políticos* encontrarán una guía detallada de cómo instalar estos programas. Tanto R como R Studio son de acceso libre y gratuito.

Contenidos y Cronograma

Núcleo Temático 1: Conociendo nuestros datos

Base de Datos: Cómo se organiza y construye una base de datos, qué va en las filas, qué va en las columnas, tipos de variables.

Variables: Cuáles son dependientes, características. Cuáles son independientes, características. Posibles relaciones entre variables. Qué análisis se puede hacer según las características de mi variable dependiente. Introducción al análisis multivariado.

Introducción a R: transformación de variables, cambio del contenido de las variables y creación de nuevas variables. Etiquetado. Filtros. Descriptivos. Frecuencias y Tablas. Resúmenes de variables (summarize).

Núcleo temático 2: Análisis bivariado

Medidas descriptivas. Medidas de tendencia central. Medidas de dispersión. Coeficientes de correlación. Tablas ANOVA. Cómo construir Tablas de contingencia (Regla de Zeisel, qué va en las filas y en las columnas, cómo se leen los porcentajes, estadísticos).

Núcleo temático 3: Análisis multivariado descriptivo I

Análisis multivariado, por qué usarlo, para qué, qué tipos de análisis multivariado hay. Análisis multivariado descriptivo y causal. Cómo medir conceptos multidimensionales. Algunas recetas de posibles análisis exploratorios multivariados: Análisis de componentes principales y Análisis Factorial

Núcleo temático 4: Análisis multivariado descriptivo II

Cómo mapear a partir de nuestros datos. Algunas recetas de posibles análisis exploratorios multivariados: Multidimensional Scaling (MDS)

Núcleo temático 5: Análisis multivariado explicativo. Análisis de regresión

Análisis de regresión. Qué debo tener en cuenta, supuestos, estadísticos de ajuste, diagnósticos, interpretación. Regresión Lineal simple (OLS).

Bibliografía y materiales de referencia:

Núcleo Temático 1: Conociendo nuestros datos

- Blalock. Hubert M (1966) *Estadística Social*. México: Fondo de Cultura Económica. Versión original en inglés (1960)

- Sharma, S. (1996) *Applied Multivariate Techniques*. New York: John Wiley.

- Urdinez, Francisco y Labrín, Andrés Cruz (eds.) (2019) Capítulo 3 y 4, en *AnalizaR Datos Políticos*; versión 0.1

Podrán acceder al libro mediante este link:

<https://arcruz0.github.io/libroadp/index.html>

Núcleo temático 2: Análisis bivariado

- Kellstedt, Paul and Guy Whitten (2009) *The Fundamentals of Political Science Research*. Cambridge University Press. Capítulos 7 y 8.

- Baranger, Denis (2009) *Construcción y análisis de datos. Introducción al uso de técnicas cuantitativas en la investigación social*. Posadas. Capítulo 4.

- Ramos González, Ramón, "Contraste de hipótesis". Materiales de cátedra, *Métodos Estadísticos y numéricos*. I.E.S. A Xunqueira I

- Dietrichson, Aleksander (2019) *Métodos Cuantitativos*.
<https://bookdown.org/dietrichson/metodos-cuantitativos/>

Núcleo temático 3: Análisis multivariado exploratorio

- Manly, Bryan F. J. (1985). *Multivariate Statistical Methods*. Capítulo 1, 5 y 6. Londres: Chapman and Hall. Sage Publications.

- Riba, Clara (2007) "Factor Analyses. Un example using STATA." *Material de Cátedra* Universitat Pompeu Fabra. Barcelona

- Urdinez, Francisco y Labrín, Andrés Cruz (eds.) (2019) "Capítulo 15", en *AnalizaR Datos Políticos*.

- Ratto, M. C. (2007) "Eliendo las reglas del juego. El nivel de competencia efectiva de las reformas electorales sub-nacionales argentinas, 1983-2005", en Juan Abal Medina (comp), *Evaluando el Desempeño Democrático de las Instituciones Políticas Argentinas*. Prometeo: Buenos Aires.

- Reynoso, D (2023) "La naturaleza de la ideología política del electorado argentino", *Revista Argentina de Ciencia Política*, Vol. I, Núm. 30, pp. 158-189 (<https://publicaciones.sociales.uba.ar/index.php/revistaargentinacienciapolitica/article/view/9008/7601>)

Núcleo temático 4: Análisis multivariado exploratorio II

- Manly, Bryan F. J. (1985). *Multivariate Statistical Methods*. Capítulo 10. Londres: Chapman and Hall. Sage Publications.

- Riba, Clara (2007) "Multidimensional Scaling. An example using STATA. Political parties in the Spanish political space." *Material de Càtedra* Universitat Pompeu Fabra. Barcelona

- Ratto, M. C. (2018) "Ideología, votantes y partidos: el espacio ideológico de la competencia partidaria en la Argentina". *Question*, Vol. 1 Núm. 60 (2018): Primavera (octubre-diciembre). doi:<https://doi.org/10.24215/16696581e105>

Núcleo temático 5: Análisis de regresión

- Kellstedt, Paul and Guy Whitten (2009) *The Fundamentals of Political Science Research*. Cambridge University Press. Capítulos 9, 10 y 11.
- King, Gary, Michael Tomz, and Jason Wittenberg (2000) "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science*, 44:347-61.
- Lewis-Beck, M. (1980). *Applied Regression: An Introduction*. SAGE Publications.
- Fynn, Inés y Nocetto, Lihuen (2019) "Capítulo 5: Modelos Lineales", en Urdinez, Francisco y Labrín, Andrés Cruz (eds.) (2019) *AnalizaR Datos Políticos*.
- Reynoso, Diego (2021) "Explicando las actitudes hacia el aborto: religiosidad, ideología, región y educación", *Revista SAAP*, Vol. 15, N° 1, mayo 2021, 13-41.