

ITAM
Departamento de Ciencia Política

Metodología Política Avanzada

Primavera 2022
Lunes 18:00-21:00
Río Hondo, Salón CC102

Adrián Lucardi
adrian.lucardi@itam.mx

Oficina: Lunes y Miércoles 14:00-15:00 o previa cita vía Canvas

En este curso veremos los rudimentos del análisis cuantitativo en ciencias sociales, poniendo especial énfasis en (i) la *manipulación* y *visualización* de datos; (ii) la *especificación* de modelos estadísticos; y (iii) la *interpretación* de dichos modelos.

El curso parte de la idea de que la estadística no es una técnica ni una metodología sino un lenguaje, y la única forma de aprender un lenguaje es practicándolo. Por lo tanto, para aprobar el curso deberás completar numerosos ejercicios prácticos *fuera del horario de clases*. Tanto el material teórico como los ejercicios de manipulación y análisis de datos son acumulativos, por lo que es indispensable mantenerse al día con las tareas.

Para tomar el curso no necesitas estar familiarizado con ningún lenguaje de programación. Sin embargo, en la presentación del material teórico daré por sentado que conoces los rudimentos de la teoría de probabilidades y el modelo de regresión por mínimos cuadrados ordinarios. A lo largo del curso discutiremos cuándo es posible interpretar los resultados de los modelos en términos causales, pero el foco estará puesto en la especificación e interpretación de modelos estadísticos, no en la inferencia causal como tal.

Calificación

- **Participación en clase (11%).** Debes leer la bibliografía obligatoria y participar activamente en las discusiones, sea personalmente o en el chat.
- **Tareas (36%).** A lo largo de la clase habrá un total de ocho tareas, pero solo las *seis mejores* contarán para la calificación final. Puedes responder las tareas de forma individual o en conjunto con *un/a único/a compañero/a*, cuyo nombre deberá estar claramente especificado en la hoja de respuesta. La fecha límite para entregar las tareas es la indicada en este temario, *antes* del comienzo de la clase. Cada tarea será calificada como (i) entregada completa y aprobada (6% de la calificación final); (ii) entregada incompleta y/o deficiente (3%); o (iii) no entregada (0%).
- **Parcial (20%).** El lunes 25 de abril habrá un examen parcial en horario de clase, a libro abierto pero a computadora cerrada.

- **Trabajo final (33%).** Debes replicar los resultados –tablas y gráficos– de un artículo publicado en una revista académica de ciencias sociales. La calificación del trabajo final se distribuirá de la siguiente manera:
 - *Selección.* A más tardar el jueves 28 de abril deberás subir a Canvas una nota mencionando (i) la referencia del artículo que quieres replicar; (ii) las tablas y/o figuras que te propones replicar, así como los análisis adicional(es) que te propones llevar a cabo; y (iii) adjuntando la base de datos correspondiente y/o la liga donde descargar los mismos. Puedes proponer más de una opción si lo deseas. Enviar tu selección a tiempo contará por el 3% de la calificación final.

La replicación *puede* ser en equipo, sujeta a las siguientes restricciones:

 - i. Ningún equipo puede tener más de dos integrantes;
 - ii. Debe haber al menos una propuesta de replicación por miembro; y
 - iii. Las replicaciones grupales deben ser más extensas –en el sentido de cubrir más figuras y tablas del artículo original– que las individuales.
 - *Replicación.* En fecha a definir durante el período de finales, cada estudiante o equipo debe enviar (i) un reporte detallando los resultados de la replicación; y (ii) el código para reproducir los mismos. El reporte contará el 30% de la calificación final.

Lecturas

El foco de la clase estará puesto en cuestiones prácticas, no teóricas. Si deseas profundizar (algunas) cuestiones teóricas, puedes consultar los siguientes textos, ordenados de menor a mayor grado de dificultad:

- Ethan Bueno de Mesquita and Anthony Fowler. *Thinking Clearly with Data: A Guide to Quantitative Reasoning and Analysis*. Princeton University Press, Princeton, 2021.
- Joshua D. Angrist and Jörn-Steffen Pischke. *Mastering 'Metrics. The Path from Cause to Effect*. Princeton University Press, Princeton, 2015.
- Kosuke Imai. *Quantitative Social Science: An Introduction*. Princeton University Press, Princeton and Oxford, 2018.
- Joshua Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics. An Empiricist's Companion*. Princeton University Press, Princeton, 2009.
- Jeffrey M. Wooldridge. *Introductory Econometrics: A Modern Approach*. 5th edition, 2012.

Todas las lecturas se encuentran disponibles en Canvas.

Software

Junto con las tareas y el trabajo final, deberás adjuntar el correspondiente código de R ó Stata. No aceptaré otros lenguajes de programación. Nota que todos los ejemplos de

código que distribuiré a lo largo del curso serán en R, por lo que a menos que ya estés familiarizado con Stata, te recomiendo hacer tus trabajos en R.

Puedes descargar R de manera gratuita en <https://cran.itam.mx/>. Te recomiendo usarlo en conjunto con RStudio (<https://www.rstudio.com/>). Si quieres usar Stata, el ITAM tiene licencias para profesores y alumnos; puedes descargarlo desde [Comunidad](#).

La manera más efectiva de aprender a programar es buscando código de otros autores en internet y adaptándolo a tus propias necesidades. Si eres nuevo en R, los siguientes recursos te pueden resultar útiles:

- Kelsey Moty, *workshop* de verano de NYU: <https://nyu-cdsc.github.io/learningr/>.
- Francisco Urdinez and Andrés Cruz Labrín, editors. *AnalizaR Datos Políticos*. 2020: <https://arcruz0.github.io/libroadp/>. Caps. 1-4, 11 y 16 únicamente.
- Hadley Wickham and Garrett Grolemund. *R for Data Science. Visualize, Model, Transform, Tidy and Import Data*. O'Reilly, Sebastopol (CA), 2017: <https://r4ds.had.co.nz/>.
- Si quieres entregar tus tareas en Markdown (opcional): Yihui Xie, Christophe Dervieux, and Emily Riederer. *R Markdown Cookbook*. Chapman and Hall/CRC, Boca Raton, 2020: <https://bookdown.org/yihui/rmarkdown-cookbook/>.

Integridad académica

Según el artículo 5 bis del Reglamento de Titulación del ITAM, “la reproducción parcial o total de textos o productos ajenos, de autor conocido o anónimo, publicados por cualquier medio o inéditos, sin entrecomillar los pasajes o elementos reproducidos o sin hacer la referencia bibliográfica pertinente,” constituye plagio o fraude académico. Las citas a otros autores son bienvenidas, pero deben ser reconocidas como tales. El uso de fragmentos y/o traducciones literales de textos ajenos sin la debida atribución será causal para reprobación de la materia y merecerá la amonestación oficial al alumno por parte de la Dirección Escolar del Instituto.

Calendario

Sesión 1 (Enero/24). Introducción. Manipulación de datos (1)

- Presentación del curso.
- *Teoría*: Tipos de datos
 - *Lectura optativa*: Wooldridge, sec. 1.3.
- *Práctica*: introducción a R. Operadores condicionales; selección de subconjuntos; combinación de bases de datos; estadísticas descriptivas
 - Kelsey Moty, *workshop* NYU: <https://nyu-cdsc.github.io/learningr/>. Lecturas y ejercicios de “Week 1: Intro to R” (“Part 1” y “More about variables”).

- Hadley Wickham and Garrett Grolemund. *R for Data Science. Visualize, Model, Transform, Tidy and Import Data*. O’Reilly, Sebastopol (CA), 2017: “[Data transformation](#)” (excepto la parte correspondiente a `ggplot2`).
- *Opcional*: Hadley Wickham and Garrett Grolemund. *R for Data Science. Visualize, Model, Transform, Tidy and Import Data*. O’Reilly, Sebastopol (CA), 2017: “[Data import](#)”.
- Consigna para la *Tarea #1*.
- *Datos*:
 - Ideología en las votaciones de la ONU: Michael A. Bailey, Anton Strezhnev, and Erik Voeten. Estimating Dynamic State Preferences from United Nations Voting Data. *Journal of Conflict Resolution*, 61(2):430–456, 2017.
 - Ideología presidencial en América Latina.
 - Ministros alrededor del mundo: Jacob Nyrupe and Stuart Bramwell. Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4):1366–1374, 2020.

Sesión 2 (Enero/31). Manipulación de datos (II): Agrupación por categorías

- *Práctica*: datos agrupados por categorías; *leads* y *lags*
 - Roger D. Peng. *R Programming for Data Science*, 2016, cap. 2: <https://bookdown.org/rdpeng/rprogdatascience/managing-data-frames-with-the-dplyr-package.html>.
 - Kelsey Moty, *workshop NYU*: <https://nyu-cdsc.github.io/learningr/>. Lecturas y ejercicios de “Week 4” (“Learning the main 6 dplyr verbs,” esp. el código correspondiente a la base de datos `disgust`).
 - Hadley Wickham and Garrett Grolemund. *R for Data Science. Visualize, Model, Transform, Tidy and Import Data*. O’Reilly, Sebastopol (CA), 2017: “[Relational data](#)”.
 - *Opcional*: Hadley Wickham and Garrett Grolemund. *R for Data Science. Visualize, Model, Transform, Tidy and Import Data*. O’Reilly, Sebastopol (CA), 2017: “[Tidy data](#)”.
 - *Optional*: Kelsey Moty, *workshop NYU*: <https://nyu-cdsc.github.io/learningr/>. Lecturas y ejercicios de “Week 4” (“Describing data”);
 - Entrega y revisión de la *Tarea #1*.
 - Consigna para la *Tarea #2*.
- *Datos*:
 - Regímenes políticos y desempeño económico, 1946-2010.
 - Índice de Gini: `swiid` (<https://fsolt.org/swiid/>). Frederick Solt. The Standardized World Income Inequality Database. *Social Science Quarterly*, 97(5):1267–1281, 2016.
 - Ideología en las votaciones de la ONU: Michael A. Bailey, Anton Strezhnev, and Erik Voeten. Estimating Dynamic State Preferences from United Nations Voting Data. *Journal of Conflict Resolution*, 61(2):430–456, 2017.
 - Ideología presidencial en América Latina.

- Ministros alrededor del mundo: Jacob Nyrop and Stuart Bramwell. Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4):1366–1374, 2020.
- LAPOP: Promedios por países.

Febrero/7. Día de la Bandera

No hay clase.

Sesión 3 (Febrero/14). Manipulación de datos (III): Visualización

- *Teoría*: buenos y malos gráficos
 - *Lectura optativa*: Kieran Healy. *Data Visualization: A Practical Introduction*. Princeton University Press, Princeton and Oxford, 2018, cap. 1, en <http://socviz.co/>.
 - *Lectura optativa*: Sarah Leo, “Mistakes, We’ve Drawn a Few”.
- *Práctica*: visualización de datos; la lógica de `ggplot2`.
 - Kieran Healy. *Data Visualization: A Practical Introduction*. Princeton University Press, Princeton and Oxford, 2018, caps. 3-4, en <http://socviz.co/>.
 - *Opcional*: Kieran Healy. *Data Visualization: A Practical Introduction*. Princeton University Press, Princeton and Oxford, 2018, cap. 5, en <http://socviz.co/>.
 - *Opcional*: Hadley Wickham and Garrett Grolemund. *R for Data Science. Visualize, Model, Transform, Tidy and Import Data*. O’Reilly, Sebastopol (CA), 2017: “Data visualization”.
 - Entrega y revisión de la *Tarea #2*.
 - Consigna para la *Tarea #3*.
- *Datos*:
 - Número de hijas y liberalismo en el Congreso norteamericano: Ebonya L. Washington. Female Socialization: How Daughters Affect Their Legislator Fathers’ Voting on Women’s Issues. *American Economic Review*, 98(1):311–332, 2008.
 - Ideología en las votaciones de la ONU: Michael A. Bailey, Anton Strezhnev, and Erik Voeten. Estimating Dynamic State Preferences from United Nations Voting Data. *Journal of Conflict Resolution*, 61(2):430–456, 2017.
 - Ideología presidencial en América Latina.
 - Ministros alrededor del mundo: Jacob Nyrop and Stuart Bramwell. Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4):1366–1374, 2020.
 - LAPOP: Promedios por países.

Sesión 4 (Febrero/21). Manipulación de datos (IV): Visualización y agrupación por categorías (cont.)

- *Práctica*: agrupación por categorías + visualización de datos.
 - Continuación de la *Tarea #3*.

- *Datos:*
 - Número de hijas y liberalismo en el Congreso norteamericano: Ebonya L. Washington. Female Socialization: How Daughters Affect Their Legislator Fathers' Voting on Women's Issues. *American Economic Review*, 98(1):311–332, 2008.
 - Ideología en las votaciones de la ONU: Michael A. Bailey, Anton Strezhnev, and Erik Voeten. Estimating Dynamic State Preferences from United Nations Voting Data. *Journal of Conflict Resolution*, 61(2):430–456, 2017.
 - Ideología presidencial en América Latina.
 - Ministros alrededor del mundo: Jacob Nyrup and Stuart Bramwell. Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4):1366–1374, 2020.
 - LAPOP: Promedios por países.

Sesión 5 (Febrero/28). Manipulación de datos (v): Sistemas de Información Geográfica (GIS)

- *Teoría:* Estructura de datos con GIS
 - *Lectura optativa:* Kieran Healy. *Data Visualization: A Practical Introduction*. Princeton University Press, Princeton and Oxford, 2018, cap. 7.
 - *Lectura optativa:* Kosuke Imai. *Quantitative Social Science: An Introduction*. Princeton University Press, Princeton and Oxford, 2018, sec. 5.3.
- *Práctica:* coloreando y manipulando mapas
 - Mel Moreno y Mathieu Basille, “Drawing Beautiful Maps Programmatically with R, sf and ggplot2,” [Parte 1](#) y [Parte 2](#).
 - Entrega y revisión de la *Tarea #3*.
 - Consigna para la *Tarea #4*.
- *Datos:*
 - Resultados electorales en México (2018).
 - Votación del matrimonio igualitario (2010) en Argentina.

Sesión 6 (Marzo/7). Regresión

- *Teoría:* regresión simple y múltiple
 - *Lectura optativa:* Wooldridge, secs. 2.1-2.4, 3.1-3.2 y 7.2-7.3.
- *Práctica:* interpretación de coeficientes
- *Datos:*
 - El Cuarteto de Anscombe.
 - Número de hijas y liberalismo en el Congreso norteamericano: Ebonya L. Washington. Female Socialization: How Daughters Affect Their Legislator Fathers' Voting on Women's Issues. *American Economic Review*, 98(1):311–332, 2008.
 - Gastos de Campaña y Porcentaje de Votos en USA (2018): Ethan Bueno de Mesquita and Anthony Fowler. *Thinking Clearly with Data: A Guide to Quantitative Reasoning and Analysis*. Princeton University Press, Princeton, 2021.

Sesión 7 (Marzo/14). Inferencia

- *Teoría*: significancia estadística + heteroscedasticidad
 - Ethan Bueno de Mesquita and Anthony Fowler. *Thinking Clearly with Data: A Guide to Quantitative Reasoning and Analysis*. Princeton University Press, Princeton, 2021, caps. 6-7.
 - *Lectura optativa*: Wooldridge, secs. 4.1-4.3 y 8.1-8.2.
 - *Lectura optativa*: Jeff Gill. The Insignificance of Null Hypothesis Significance Testing. *Political Research Quarterly*, 52(3):647–674, 1999.
 - *Lectura optativa*: Carlisle Rainey. Arguing for a Negligible Effect. *American Journal of Political Science*, 58(4):1083–1091, 2014.
 - *Lectura optativa*: Donald P. Green, Mary C. McGrath, and Peter M. Aronow. Field Experiments and the Study of Voter Turnout. *Journal of Elections, Public Opinion and Parties*, 23(1):27–48, 2013.
- *Práctica*: ajustando los errores estándar
 - Entrega y revisión de la *Tarea #4*.
 - Consigna para la *Tarea #5*.
- *Datos*:
 - Número de hijas y liberalismo en el Congreso norteamericano: Ebonya L. Washington. Female Socialization: How Daughters Affect Their Legislator Fathers' Voting on Women's Issues. *American Economic Review*, 98(1):311–332, 2008.
 - LAPOP: opiniones sobre liderazgo femenino en América Latina.
 - Ministros alrededor del mundo: Jacob Nyrup and Stuart Bramwell. Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4):1366–1374, 2020.

Marzo/21. Natalicio de Benito Juárez

No hay clase.

Sesión 8 (Marzo/28). Interacciones

- *Teoría*: interpretación
 - *Lectura optativa*: Thomas Brambor, William R. Clark, and Matt Golder. Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis*, 14(1):63–82, 2006.
- *Práctica*: incluyendo interacciones
 - Entrega y discusión de la *Tarea #5*.
- *Datos*:
 - Votación del matrimonio igualitario (2010) en Argentina.
 - Número de hijas y liberalismo en el Congreso norteamericano: Ebonya L. Washington. Female Socialization: How Daughters Affect Their Legislator Fathers' Voting on Women's Issues. *American Economic Review*, 98(1):311–332, 2008.
 - Regímenes políticos y desempeño económico, 1946-2010.

- Ministros alrededor del mundo: Jacob Nyrup and Stuart Bramwell. Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4):1366–1374, 2020.

Sesión 9 (Abril/4). Variación *within* vs. *between*

- *Teoría*: variación “within” y “between”
 - Jonathan Mummolo and Erik Peterson. Improving the Interpretation of Fixed Effects Regression Results. *Political Science Research and Methods*, 6(4):829–835, 2018.
 - *Lectura optativa*: Chad Hazlett and Leonard Wainstein. Understanding, Choosing, and Unifying Multilevel and Fixed Effect Approaches. *Political Analysis*, 30(1):46–65, 2022.
 - *Lectura optativa*: Andrew Bell and Kelvyn Jones. Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data. *Political Science Research and Methods*, 3(1):133–153, 2015.
- *Práctica*: corriendo modelos *pooled*, *within* y *between*; visualizando datos *pooled*, *within* y *between*
 - Consigna para la Tarea #6.
- *Datos*:
 - Votación del matrimonio igualitario (2010) en Argentina.
 - Ideología en las votaciones de la ONU: Michael A. Bailey, Anton Strezhnev, and Erik Voeten. Estimating Dynamic State Preferences from United Nations Voting Data. *Journal of Conflict Resolution*, 61(2):430–456, 2017.
 - Ideología presidencial en América Latina.
 - LAPOP: opiniones sobre liderazgo femenino en América Latina.
 - Ministros alrededor del mundo: Jacob Nyrup and Stuart Bramwell. Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4):1366–1374, 2020.

Abril/11. Semana Santa

No hay clase.

Sesión 10 (Abril/18). Variables dependientes discretas

- *Teoría*: *logit*, *probit* y el modelo de probabilidad lineal
 - Walter Sosa Escudero, “Regresión logística ¿Para todos, algunos o nadie?”
 - *Lectura optativa*: Wooldridge, secs. 7.5 y 17.1.
- *Práctica*: convirtiendo coeficientes de modelos *logit/probit* en probabilidades
 - Entrega y discusión de la Tarea #6.
- *Datos*:
 - Votación del matrimonio igualitario (2010) en Argentina.

- Regímenes políticos y desempeño económico, 1960-2015.
- LAPOP: opiniones sobre liderazgo femenino en América Latina.
- Ministros alrededor del mundo: Jacob Nyrup and Stuart Bramwell. Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4):1366–1374, 2020.

Sesión 11 (Abril/25). Examen parcial

- Examen parcial. En horario de clase.

Sesión 12 (Mayo/2). Datos de panel + Diferencia-en-Diferencias (I): Introducción

- *Teoría*: diferencia-en-diferencias
 - Paul J. Gertler, Sebastián Martínez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch. *Impact Evaluation in Practice [Second Edition]*. The World Bank-IDB, Washington DC, 2nd edition, 2016, cap. 7.
 - *Lectura optativa*: Joshua D. Angrist and Jörn-Steffen Pischke. *Mastering 'Metrics. The Path from Cause to Effect*. Princeton University Press, Princeton, 2015, cap. 5.
- *Práctica*: DiD con *pooled cross sections* y paneles
 - Consigna para la *Tarea #7*.
- *Datos*:
 - Terrorismo en el País Vasco: Alberto Abadie and Javier Gardeazabal. The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review*, 93(1):113–132, 2003.
 - Ideología en las votaciones de la ONU: Michael A. Bailey, Anton Strezhnev, and Erik Voeten. Estimating Dynamic State Preferences from United Nations Voting Data. *Journal of Conflict Resolution*, 61(2):430–456, 2017.
 - Revoluciones sociales: Jean Lachapelle, Steven Levitsky, Lucan A. Way, and Adam E. Casey. Social Revolution and Authoritarian Durability. *World Politics*, 72(4):557–600, 2020.
 - LAPOP: opiniones sobre liderazgo femenino en América Latina.
 - Ministros alrededor del mundo: Jacob Nyrup and Stuart Bramwell. Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4):1366–1374, 2020.

Sesión 13 (Mayo/9). Diferencia-en-Diferencias (II): Problemas y Extensiones

- *Teoría*: diferencia-en-diferencias
 - *Lectura optativa*: Andrew C. Baker, David F. Larcker, and Charles C. Y. Wang. How Much Should We Trust Staggered Difference-In-Differences Estimates? *Journal of Financial Economics*, forthcoming.

- *Lectura optativa*: Andrew Goodman-Bacon. Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics*, 225(2):254–277, 2021 + FAQ.
- *Práctica*: DiD con *pooled cross sections* y paneles. *Event studies*
 - Continuación de la *Tarea #7*.
- *Datos*:
 - Terrorismo en el País Vasco: Alberto Abadie and Javier Gardeazabal. The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review*, 93(1):113–132, 2003.
 - Regímenes políticos y desempeño económico, 1946-2010.
 - Ideología en las votaciones de la ONU: Michael A. Bailey, Anton Strezhnev, and Erik Voeten. Estimating Dynamic State Preferences from United Nations Voting Data. *Journal of Conflict Resolution*, 61(2):430–456, 2017.
 - Ideología presidencial en América Latina.
 - LAPOP: opiniones sobre liderazgo femenino en América Latina.
 - Ministros alrededor del mundo: Jacob Nyrop and Stuart Bramwell. Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4):1366–1374, 2020.

Sesión 14 (Mayo/16). Regresión discontinua

- *Teoría*: El modelo de regresión discontinua
 - *Lectura optativa*: Matías D. Cattaneo, Rocío Titiunik, and Gonzalo Vázquez-Bare. The Regression Discontinuity Design. In Luigi Curini and Robert J. Franzese, editors, *The SAGE Handbook of Research Methods in Political Science and International Relations [Two Volume Set]*, pages 835–857. Sage, 2020.
- *Práctica*: visualización y estimación de regresiones discontinuas
 - Entrega y revisión de la *Tarea #7*.
 - Consigna para la *Tarea #8*.
- *Datos*:
 - Resultados electorales en México.
 - ¿El Distrito o el Partido?: Ethan Bueno de Mesquita and Anthony Fowler. *Thinking Clearly with Data: A Guide to Quantitative Reasoning and Analysis*. Princeton University Press, Princeton, 2021.

Sesión 15 (Mayo/23). Repaso

- Entrega y revisión de la *Tarea #8*.
- Repaso y discusión del trabajo final. Asistencia optativa.

Fecha a definir. Trabajo final

Durante segunda semana de finales.
Subir a Canvas antes de las 10:00am.