

ITAM
Departamento de Ciencia Política

Metodología Política Avanzada

Otoño 2021
Lunes 17:30-20:30
Río Hondo, Salón CC102 (cruzamos los dedos)

Adrián Lucardi
adrian.lucardi@itam.mx
Oficina: Lunes y Miércoles 12:00-13:00 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 (10%).** Debes leer la bibliografía obligatoria y participar activamente en las discusiones, sea personalmente o en el chat.
- **Tareas (39%).** 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 (6.5% de la calificación final); (ii) entregada incompleta y/o deficiente (3%); o (iii) no entregada (0%).
- **Parcial (18%).** El lunes 11 de octubre habrá un examen parcial en horario de clase.

- **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 octubre deberás subir a Canvas una nota (i) indicando la referencia del artículo que quieres replicar; (ii) detallando qué tablas y/o figuras te propones replicar, así como qué análisis adicional(es) 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. El proceso de selección es individual: cada miembro debe proponer al menos un artículo; y
 - iii. Las replications 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:

- 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.
- Fred L. Ramsey and Daniel W. Schafer. *The Statistical Sleuth. A Course in Methods of Data Analysis [Second Edition].* Brooks/Cole, Belmont, CA, 2nd edition, 2002.
- Peter Kennedy. *A Guide to Econometrics.* Wiley-Blackwell, Malden, MA, 6th edition, 2008.
- 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 código en otros lenguajes de programación. No obstante, 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 reprobar la materia y merecerá la amonestación oficial al alumno por parte de la Dirección Escolar del Instituto.

Calendario

Sesión 1 (Agosto/9). Introducción

- Presentación del curso.
- *Teoría*: Tipos de datos
 - *Lectura optativa*: Wooldridge, sec. 1.3.
- *Práctica*: introducción a R
 - Kelsey Moty, *workshop* NYU: <https://nyu-cdsc.github.io/learningr/>. Lecturas y ejercicios de “Week 1: Intro to R” (“Part 1,” “Part 2” y “More about packages”).
 - Kieran Healy. *Data Visualization: A Practical Introduction*. Princeton University Press, Princeton and Oxford, 2018, secs. 2.2-2.7: <http://socviz.co/gettingstarted.html>
 - *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 Nyrup and Stuart Bramwell. Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4):1366–1374, 2020.
 - *Global Leadership Project* (GLC): John Gerring, Erzen Oncel, Kevin Morrison, and Daniel Pemstein. Who Rules the World? A Portrait of the Global Leadership Class. *Perspectives on Politics*, 17(4):1079–1097, 2019.

Sesión 2 (Agosto/16). Manipulación de datos (I)

- *Práctica*: operadores condicionales; selección de subconjuntos; combinación de bases de datos; estadísticas descriptivas por grupos
 - Kelsey Moty, *workshop* NYU: <https://nyu-cdsc.github.io/learningr/>. Lecturas y ejercicios de “Week 1” (“More about variables” y “More about vectors”); “Week 4” (“Describing data”); y “Week 5” (“Joining data using dplyr’s join verbs”).
 - 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) y “Relational data”.
 - Continuación de 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 Nyrop and Stuart Bramwell. Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4):1366–1374, 2020.
- *Global Leadership Project* (GLC): John Gerring, Erzen Oncel, Kevin Morrison, and Daniel Pemstein. Who Rules the World? A Portrait of the Global Leadership Class. *Perspectives on Politics*, 17(4):1079–1097, 2019.

Sesión 3 (Agosto/23). 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 3” (“Manipulating your data using `tidyr`”) y “Week 4” (“Learning the main 6 `dplyr` verbs,” esp. el código correspondiente a la base de datos `disgust`).
 - *Opcional*: Hadley Wickham and Garrett Grolemund. *R for Data Science. Visualize, Model, Transform, Tidy and Import Data*. O’Reilly, Sebastopol (CA), 2017: “Tidy 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.

Sesión 4 (Agosto/30). 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.

Sesión 5 (Septiembre/6). 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
 - Continuación de la *Tarea #3*.
- *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.
 - LAPOP: opiniones sobre liderazgo femenino en América Latina.

Sesión 6 (Septiembre/13). Manipulación de datos (IV): 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.
 - *Global Leadership Project (GLC)*: John Gerring, Erzen Oncel, Kevin Morrison, and Daniel Pemstein. Who Rules the World? A Portrait of the Global Leadership Class. *Perspectives on Politics*, 17(4):1079–1097, 2019.

Sesión 7 (Septiembre/20). Inferencia

- *Teoría*: significancia estadística + heteroscedasticidad
 - Claudia Goldin and Cecilia Rouse. Orchestrating Impartiality: The Impact of “Blind” Auditions on Female Musicians. *American Economic Review*, 90(4):58, 2000.
 - Andrew Gelman, “[Did Blind Orchestra Auditions Really Benefit Women?](#)”
 - *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.
- *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 Nyrop and Stuart Bramwell. Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4):1366–1374, 2020.

Sesión 8 (Septiembre/27). Datos agrupados: 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*, forthcoming.
- *Práctica*: corriendo modelos con efectos fijos vs. aleatorios; visualización de datos de panel

- *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.

Sesión 9 (Octubre/4). 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*.
 - Consigna para la *Tarea #6*.
- *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 10 (Octubre/11). Examen parcial

- Examen parcial. En horario de clase.

Sesión 11 (Octubre/18). 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.
- *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.
- Ideología presidencial en América Latina.
- 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 12 (Octubre/25). Diferencia-en-Diferencias (II): Problemas y Extensiones

- *Teoría*: diferencia-en-diferencias
 - *Lectura optativa*: Andrew Baker, David F. Larcker, and Charles C. Y. Wang. How Much Should We Trust Staggered Difference-In-Differences Estimates? SSRN Scholarly Paper, SSRN, 2021. How Much Should We Trust Staggered DiD?
 - *Lectura optativa*: Andrew Goodman-Bacon. Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics*, forthcoming + FAQ.
- *Práctica*: DiD con *pooled cross sections* y paneles. *Event studies*
 - Entrega y discusión de la *Tarea #6*.
 - 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.
 - 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 Nyrup and Stuart Bramwell. Who Governs? A New Global Dataset on Members of Cabinets. *American Political Science Review*, 114(4):1366–1374, 2020.

Noviembre/1. Día de Muertos

- No hay clase.

Sesión 13 (Noviembre/8). 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 #7*.
 - Consigna para la *Tarea #8*.
- *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.

Noviembre/15. Día de la Revolución

No hay clase.

Sesión 14 (Noviembre/22). 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 #8*.
- *Datos*:
 - Resultados electorales en México.

Sesión 15 (Noviembre/29). Repaso

- 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.