

# Métodos cuantitativos básicos

Sebastián Garrido de Sierra

Otoño 2019

H. de clases: Lun. y Mie. 10:10-12:10 hrs.  
Salón: Teatro de Decisiones (Piso 3 del LNPP)  
[sebastian.garrido@cide.edu](mailto:sebastian.garrido@cide.edu)

H. de oficina: Lun. 15-16:00 hrs.  
Oficina: Piso 2 del LNPP  
[segasi.mx](mailto:segasi.mx)

## ¿De qué va este curso?

Éste es el primer curso de la secuencia de métodos de la Licenciatura en Ciencia Política y Relaciones Internacionales del CIDE.

Su principal objetivo es dotarte de sólidas bases conceptuales y técnicas para que puedas transformar *datos* en *información* y *conocimiento* adecuado para el consumo humano.

Este proceso está compuesto de varias etapas que pueden resumirse gráficamente de la siguiente forma:

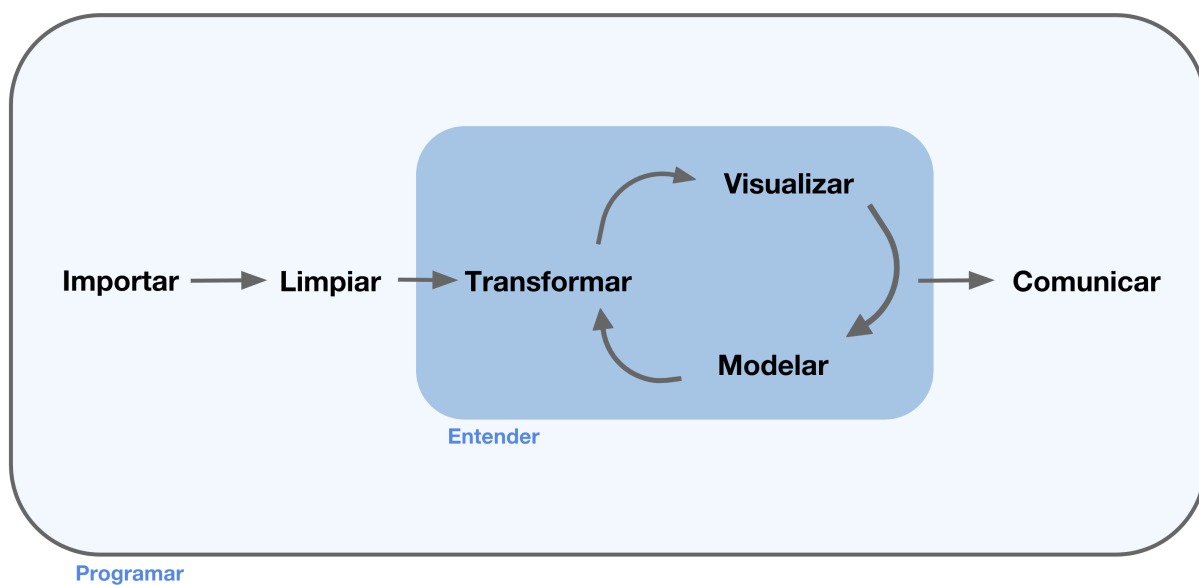


Figura 1: Basado en [Grolemund y Wickham, 2017](#)

Para lograr el objetivo recién planteado, durante las próximas 15 semanas (y cachito) cubriremos diversos temas relacionados con:<sup>1</sup>

1. Manejo de bases de datos (e.g., unión y limpieza de bases de datos)
2. Visualización de datos (e.g., cómo hacer gráficas y mapas eficientes + efectivos)
3. Estadística descriptiva (e.g., medidas de tendencia central y dispersión)
4. Inferencia y probabilidad (e.g., distribuciones de probabilidad, pruebas de hipótesis)
5. Programación (casi desde el comienzo aprenderán a usar R y RStudio)

## Formato del curso

Nos reuniremos dos veces a la semana, lunes y miércoles, de 10:10 a 12:10 hrs. en el Teatro de Decisiones del LNPP, ubicado en el piso 3 del LNPP (mismo edificio que la Biblioteca).

Gracias a la generosidad de Lila Petersen, estudiante del Doctorado en Política Pública, y Pablo Parás, estudiante de LCPRI, además tendrán la posibilidad de asistir **voluntariamente** a un laboratorio de 1:30 hr. a la semana. Ustedes definirán con Lila y Pablo cuál es el mejor día/hora para el lab.

## Evaluación

Su calificación final dependerá de su desempeño en las siguientes actividades:

Actividad	Porcentaje
Tareas	20 %
Primer examen parcial	20 %
Segundo examen parcial	20 %
Tercer examen parcial	20 %
Cuarto examen parcial	20 %

## Lecturas

Una muy buena parte de las lecturas que harán este semestre corresponden a capítulos o secciones de los siguientes cuatro libros, todos disponibles en línea de forma **gratuita**:

- [Diez, David M, Christopher D. Barr, and Mine Cetinkaya-Rundel \(2017\). \*OpenIntro Statistics\*. OpenIntro.\)](#)

<sup>1</sup>Si te interesa leer un poco sobre la motivación detrás de esta combinación de temas, te recomiendo leer los siguientes textos de [Nolan y Lang](#), [Cobb](#), [Cobb 1](#) y [American Statistical Association](#)).

- Grolemund, Garrett and Hadley Wickham (2017). *R for Data Science*. Canada: O'Reilly.)
- Ismay, Chester and Albert Y. Kim (2018). *Modern Drive. An Introduction to Statistical and Data Sciences via R.*)
- Phillips, Nathaniel D. (2018). *YaRrr! The Pirate's Guide to R.*)

Además, al comienzo del semestre les compartiré un folder en línea con el resto de las lecturas.

## Calendario

El curso consta de 29 sesiones (28 de clases y tres para exámenes), cada una de 2 hrs. A continuación detallo los temas que planeo cubrir en cada una.<sup>2</sup>

## Introducción

### Semana 01 - 08/19: ¿Por qué están aquí?

- Cap. 2, pp. 47-73 - Silver, Nate (2002). *The Signal and the Noise: Why So Many Predictions Fail-but Some Don't*. Penguin Press.
- Tversky, Amos and Daniel Kahneman (1974). "Judgment under Uncertainty: Heuristics and Biases". In: *Science* 185.4157, pp.~1124-1131.)

### Semana 01 - 08/21: ¿Cómo (aspiramos a) explicar lo que nos rodea?

- Capítulo 1 de Elster, Jon (2007). *Explaining Social Behavior. More Nuts and Bolts for the Social Science*. EEUU: Cambridge University Press.

## R, RStudio y Datos

### Semana 02 - 08/26: R y RStudio

- Secciones 2.2, 2.3, 4.4 y 9.3 de Phillips, Nathaniel D. (2018). *YaRrr! The Pirate's Guide to R*.
- Capítulo 2 de Ismay, Chester and Albert Y. Kim (2018). *Modern Drive. An Introduction to Statistical and Data Sciences via R*.

---

<sup>2</sup>Por favor toma en cuenta que el calendario del curso puede sufrir cambios dependiendo del progreso del grupo. En caso de que esto sea necesario, se los haré saber de forma anticipada.

## Semana 02 - 08/28: Tipos de datos, tipos de variables y tipos de estructuras de datos

- Secciones 1.2.1 y 1.2.2 de Diez, David M, Christopher D. Barr, and Mine Cetinkaya-Rundel (2017). *OpenIntro Statistics*. OpenIntro.
- Capítulo 2 de Munzner, Tamara (2014). *Visualization Analysis and Design*. Nueva York, EEUU: CRC Press.

## Semana 03 - 09/02: Tipos de variables y de estructuras de datos en R

### Tipos de variables

- Sección 4.4.1 de Phillips, Nathaniel D. (2018). *YaRrr! The Pirate's Guide to R*.
- Ejercicios en [DataCamp](#).

### Tipos de estructuras de datos

#### Vectores, Matrices y Dataframes

- Secciones 5.2, 8.1, 8.2 y 8.3 de Phillips, Nathaniel D. (2018). *YaRrr! The Pirate's Guide to R*.

#### Listas

- [Anexo](#) de Golemund, Garrett and Hadley Wickham (2017). *R for Data Science*. Canada: O'Reilly.

## Semana 03 - 09/04: ¿Cómo importar, unir y transformar la estructura de bases de datos en R?

- Capítulos 11, 12 y 13 de Golemund, Garrett and Hadley Wickham (2017). *R for Data Science*. Canada: O'Reilly.
- [Broman, Karl W. and Kara H. Woo \(2018\)](#). "Data organization in spreadsheets". In: *The American Statistician* 72.1, pp.~2-10.
- [Wickham, Hadley \(2014\)](#). "Tidy Data". In: *Journal of Statistical Software* 59.10.

## Semana 04 - 09/09: Jugando con datos: dplyr, parte 1

- Secciones 5.1, 5.2, 5.3 y 5.4 de Golemund, Garrett and Hadley Wickham (2017). *R for Data Science*. Canada: O'Reilly.

## **Semana 04 - 09/11: Jugando con datos: dplyr, parte 2**

- Secciones [5.5](#), [5.6](#) y [5.7](#) de Grolemund, Garrett and Hadley Wickham (2017). *R for Data Science*. Canada: O'Reilly.

## **Semana 05 - 09/16: No hay clase**

## **Semana 05 - 09/18: Jugando con datos: dplyr, parte 3**

## **Semana 06 - 09/23: Factores, fechas y cadenas de texto**

- Capítulos [14](#), [15](#) y [16](#) de Grolemund, Garrett and Hadley Wickham (2017). *R for Data Science*. Canada: O'Reilly.

## **Semana 06 - 09/25: Primer examen parcial**

## **Visualización de datos**

## **Semana 07 - 09/30: Marcas y canales**

- Capítulo 1 de Cairo, Alberto (2012). *The Functional Art: An Introduction to Information Graphics and Visualization*. California, EEUU: New Riders.
- Capítulo 5 de Munzner, Tamara (2014). *Visualization Analysis and Design*. Nueva York, EEUU: CRC Press.
- Capítulos 7 a 9 de Munzner, Tamara (2014). *Visualization Analysis and Design*. Nueva York, EEUU: CRC Press.

## **Semana 07 - 10/02: ggplot2, parte 1**

- Capítulo [3](#) de Grolemund, Garrett and Hadley Wickham (2017). *R for Data Science*. Canada: O'Reilly.
- Capítulo [3](#) de Ismay, Chester and Albert Y. Kim (2018). *Modern Drive. An Introduction to Statistical and Data Sciences via R*.

## **Semana 08 - 10/07: ggplot2, parte 2**

- Capítulo [28](#) de Grolemund, Garrett and Hadley Wickham (2017). *R for Data Science*. Canada: O'Reilly.

## Semana 08 - 10/09: ggplot2, parte 3

## Semana 09 - 10/14: Segundo examen parcial

# Estadística descriptiva

## Semana 09 - 10/16: Distribuciones univariadas

- Secciones 1.6.2 y 1.6.3 de Diez, David M, Christopher D. Barr, and Mine Cetinkaya-Rundel (2017). *OpenIntro Statistics*. OpenIntro.

## Semana 10 - 10/21: Valores típicos, atípicos y faltantes

- Secciones 3.1 y 3.2 de Agresti, Alan and Barbara Finlay (2009). *Statistical Methods for the Social Sciences*. 4th ed. Pearson.
- Sección 7.4 de Golemund, Garrett and Hadley Wickham (2017). *R for Data Science*. Canada: O'Reilly.
- Apéndice A de Ismay, Chester and Albert Y. Kim (2018). *Modern Drive. An Introduction to Statistical and Data Sciences via R*.

## Semana 10 - 10/23: Medidas de dispersión y posición

- Secciones 3.3 y 3.4 de Agresti, Alan and Barbara Finlay (2009). *Statistical Methods for the Social Sciences*. 4th ed. Pearson.
- Sección 1.6.4 de Diez, David M, Christopher D. Barr, and Mine Cetinkaya-Rundel (2017). *OpenIntro Statistics*. OpenIntro.
- Sección 7.3 de Golemund, Garrett and Hadley Wickham (2017). *R for Data Science*. Canada: O'Reilly.

## Semana 11 - 10/28: Covarianza y correlación

- Sección 7.5 de Golemund, Garrett and Hadley Wickham (2017). *R for Data Science*. Canada: O'Reilly.
- Sección 1.7 de Diez, David M, Christopher D. Barr, and Mine Cetinkaya-Rundel (2017). *OpenIntro Statistics*. OpenIntro.

## Semana 11 - 10/30: Análisis de clusters

Lecturas por definir

## **Semana 12 - 11/04: Tercer examen parcial**

### **Probabilidad**

#### **Semana 12 - 11/06: Frecuencias y valores esperados**

- Secciones 2.1.1 a 2.1.6 de Diez, David M, Christopher D. Barr, and Mine Cetinkaya-Rundel (2017). *OpenIntro Statistics*. OpenIntro.

#### **Semana 13 - 11/11: Tablas de contingencia, probabilidad marginal y conjunta**

Lecturas por definir

#### **Semana 13 - 11/13: Reglas de probabilidad, eventos (in)dependientes y probabilidad condicionada**

- Secciones 2.2 de Diez, David M, Christopher D. Barr, and Mine Cetinkaya-Rundel (2017). *OpenIntro Statistics*. OpenIntro.

#### **Semana 14 - 11/18: No hay clase**

#### **Semana 14 - 11/20: Probabilidad en R**

#### **Semana 15 - 11/25: Variables aleatorias y distribuciones de probabilidad**

- Secciones 2.4 y 2.5 de Diez, David M, Christopher D. Barr, and Mine Cetinkaya-Rundel (2017). *OpenIntro Statistics*. OpenIntro.

#### **Semana 15 - 11/27: Distribuciones de probabilidad discretas**

- Capítulo 3 de Diez, David M, Christopher D. Barr, and Mine Cetinkaya-Rundel (2017). *OpenIntro Statistics*. OpenIntro.
- Capítulo 4 de Agresti, Alan and Barbara Finlay (2009). *Statistical Methods for the Social Sciences*. 4th ed. Pearson.
- Capítulos 2 y 3 de Salsburg, David (2001). *The lady tasting tea: How statistics revolutionized science in the twentieth century*. Macmillan.

## **Semana 16 - 12/02: Distribuciones de probabilidad continuas**

Lecturas por definir

## **Semana 17 - 12/09: Examen final**