

# R @CPoliticaUBA

TAO

2021/03/23

## Presentación

En los últimos años se ha extendido el uso del análisis espacial, la visualización de datos y el análisis de modelos estadísticos para el estudio de los fenómenos sociales y la evaluación de políticas públicas.

La era de la “Big Data” propicia la adquisición de nuevas habilidades, que permitan a psicólogos, economistas, politólogos, sociólogos y otros profesionales desenvolverse con seguridad en este nuevo marco de las ciencias sociales computacionales.

El objetivo de este curso es que los estudiantes salgan sintiéndose lo suficientemente cómodos con R como para poder encontrar soluciones para sus necesidades de análisis particulares. Con ese fin, dedicaremos aproximadamente el primer tercio del curso a aprender a trabajar en el manipuleo de tablas con R, y así favorecer una transición amable a temas más complicados. Luego, se trabajará con algunos conceptos de base datos y se los aplicará en R mediante las librerías de tidyverse. En esta instancia se introducirá al mundo del análisis exploratorio de datos junto con ggplot2. Finalmente, se estudiarán los rudimentos de la programación en R: sus objetos, sentencias condicionales y estructuras de control, funciones, scoping, etc. y se le introducirán algunos algoritmos y problemas clásicos.

El curso no presupone que el alumno sabe programar. Por el contrario, presupone que el alumno quiere aprender a programar y para ello está dispuesto a estudiar ciertos fundamentos y pensar ciertos problemas en abstracto. El programa abarca el Data-Wrangling (o Munging), una introducción a la lógica de base de datos y la presentación de los conceptos iniciales de programación con el fin de dotar al estudiante autonomía en el aprendizaje y/o la capacidad de continuar su formación en cursos avanzados. En otras palabras, es un curso de introducción a la programación en R.

## Sobre el Docente

Tomás Olego es Licenciado en Ciencia Política de la Universidad de Buenos Aires (UBA). Ha realizado la Especialización y Maestría en Estadística Matemática de la UBA. Se ha desempeñado como Coordinador Estadístico en proyectos interdisciplinarios y actualmente trabaja como consultor estadístico / data scientist para proyectos de inteligencia de campañas electorales y políticas públicas basadas en evidencia. Es Profesor Adjunto de la materia Metodología de Análisis en Opinión Pública. Cuenta con diversos trabajos en revistas internacionales y ponencias en congresos.

*contacto:* tolego@sociales.uba.ar & @TAO\_dejing

## Supuestos del Docente

1. El alumno posee poca o ninguna experiencia con R (si ya usa R, al menos la primera parte le resultará aburrido).
2. El alumno no tiene experiencia previa en ningún lenguaje de programación (por ejemplo, C ++, Python, Java, etc.).

3. El alumno posee algunos conocimientos metodológicos - estadísticos al nivel de FSOC, y no le importaría repasar algunos temas.
4. El alumno está lo suficientemente motivado como para soportar la frustración y angustia (sensaciones inevitables que generarán las clases). Sobre todo, el alumno debe hallarse dispuesto al compromiso para con la temática y dedicarle tiempo y esfuerzo a este curso.

## Requisitos

1. Ser graduado de la carrera de Ciencia Política.
2. Disponer de notebook / pc.
3. Instalar R y RStudio (instrucciones cortesía de R-Ladies Buenos Aires): <http://goo.gl/FumTuJ>
4. Leer el capítulo 1 & 2 de Vázquez (2020)
5. Estudiar que es un CSV ( <https://www.computerhope.com/issues/ch001356.htm> & [https://en.wikipedia.org/wiki/Comma-separated\\_values](https://en.wikipedia.org/wiki/Comma-separated_values) )

## Temario

0. Introducción a R (Clase I)
1. Data-Wrangling con Tidyverse. (Clase I y II)
2. Introducción a la lógica de Base de Datos y su correlato en Tidyverse. (Clase III y IV)
3. Análisis exploratorio de datos con ggplot2 (Clase V)
4. Programación y estructuras de control (Clase VI)
5. Programación y funciones (Clase VII)
6. Algunos algoritmos y problemas clásicos (Clase VIII)

## Modalidad

A distancia, via Zoom.

## Cupo

Dado que se pretende tener contacto directo con el alumno, poder resolver sus dudas y acompañarlo durante la cursada, habrá cupo para de 15 estudiantes.

## Carga Horaria

16 HS totales, con 8 clases de 2HS cada una.

## Formato de Clases

Para cada clase se prevé una breve introducción teórica. Luego se desarrollará un taller en clase, durante el cual se pondrán en práctica las técnicas contempladas en el temario. Los participantes recibirán asignaciones grupales o individuales, siempre alentados a colaborar e intercambiar información.

## Bibliografia

Baumer, B. S., Kaplan, D. T., & Horton, N. J. (2017). Modern data science with R. CRC Press.

Burchell, J., & Vargas, M. (2016). The Hitchhiker's Guide to ggplot2 in R. LeanPub: Publish Early, Publish Often.

Peng, R. D. (2016). R programming for data science (pp. 86-181). Leanpub.

Vazquez Brust, A. H. Ciencia de Datos para Gente Sociable. [https://bitsandbricks.github.io/ciencia\\_de\\_datos\\_gente\\_sociable/](https://bitsandbricks.github.io/ciencia_de_datos_gente_sociable/)

Wickham, H., & Grolemund, G. (2016). R for data science: import, tidy, transform, visualize, and model data. " O'Reilly Media, Inc A distancia, via Zoom.