

Análisis de Datos para Políticas Públicas: Introducción a R



Instituto de Investigación
y Educación Económica

1. Fundamentación

En los últimos años el volumen de datos en el mundo ha crecido de manera exponencial. Se estima que el 90% de los datos mundiales se ha creado en los últimos dos años y se predice un crecimiento de un 40% anual. A este ritmo, probablemente se supere la capacidad de almacenamiento. Por otra parte, la mejora de capacidad de las computadoras, la automatización de procesos y la generación de múltiples datos a través de la internet de las cosas (IOT por sus siglas en inglés) presentan múltiples desafíos a diversas disciplinas, tanto desde el punto de vista teórico como desde la aplicación de métodos específicos.

A menudo, cuando se realizan análisis de grandes volúmenes de datos, uno de los lenguajes elegidos es R. R es un lenguaje de codificación de código abierto, con una gran capacidad de procesamiento de datos. Es gratuito, y posee una gran comunidad de usuarios con un repositorio extenso para poder utilizar en diferentes proyectos.

R está enfocado en el análisis estadístico, y permite realizar una amplia gama de operaciones: estudiar correlaciones, ajustar modelos, crear gráficos 3D de altísima calidad, aplicar árboles de decisión, realizar análisis clúster, análisis de componentes principales, crear redes neuronales de predicción, entre otras.

Este curso representa un primer acercamiento para quienes deben realizar el análisis de grandes volúmenes de datos en sus investigaciones o en su ámbito profesional. Por su enfoque, está dirigido a quienes deseen profundizar en el impacto de políticas públicas de distintos ámbitos (educación, salud, producción, etc.).

II.- Objetivos

Luego de realizar el curso los asistentes serán capaces de:

- Conocer el lenguaje R, software de licencia libre, como principal recurso para el análisis de datos.
- Desarrollar habilidades de programación en R para la exploración y procesamiento de bases de datos.
- Conocer R Markdown cómo herramienta de suma utilidad para generar información relevante en diversos formatos.

II.- A quién está dirigido y conocimientos previos requeridos

Por sus contenidos y enfoque, el curso está diseñado para profesionales que estén interesados en contar con nuevas herramientas que aporten para diseñar políticas públicas basado en la evidencia y para evaluar convenientemente su impacto. Es por ello, que será de interés para funcionarios públicos, investigadores, miembros de organizaciones del tercer sector, entre otros. Y, en términos más generales, a quienes deseen contar con habilidades que en los próximos años serán altamente valoradas por el mercado laboral.

No se requerirán conocimientos previos para poder inscribirse en el curso, ya que se trata de un curso introductorio.

Para un mejor aprovechamiento del curso, se requiere tener instalado R y RStudio (se proveerá de un tutorial para su instalación).

IV.- Cursado y metodología

El curso tiene una duración de 12 horas, distribuidas en 4 clases de 3 horas cada una. Se podrán agregar clases extra que no serán obligatorias, para resolver dudas que surjan durante la cursada. La modalidad de cursado será por streaming en vivo los días jueves, de 6:30 a 9:30 p.m., hora argentina, comenzando el 6 de Julio.

V.- Contenidos

Clase 1. Introducción al uso de R

Introducción a R y R Markdown. Introducción a la consola de R. Tipos de datos, comandos básicos, operadores lógicos y matemáticos. Abrir y guardar archivos. Uso de los foros y espacios de consulta.

Clase 2. Exploración y transformación de datos I

Manipulación de bases de datos. Paquete tidyverse. Descriptivos básicos.

Clase 3. Exploración y transformación de datos II

Limpieza de bases de datos con tidyverse, tips y funciones específicas.

Clase 4. Introducción a la visualización en R.

Introducción a la visualización de datos en R y al paquete ggplot2. Gramática de gráficos. Presentación de algunos de los tipos de gráficos más comunes. Funciones `geom_*()` y `aes()`.

VI. Profesora

Carla Degliantoni

Economista especializada en políticas sociales. Realizó la maestría de demografía social en la Universidad Nacional de Luján. Se desempeñó como analista de datos en AGISE del Ministerio de Salud del Gobierno de la Ciudad de Buenos Aires. Actualmente trabaja como jefa de gabinete de la Presidencia de la Auditoría General de la Ciudad de Buenos Aires

VII.- Certificado y condiciones de aprobación

El certificado de asistencia será otorgado por el Instituto de Investigación y Educación Económica (I+E). Para acceder a él, los asistentes deberán:

Asistir al menos al 80% de las horas de cursado sincrónico
Presentar y aprobar un trabajo individual de aplicación. Las pautas para la realización del trabajo serán entregadas en el transcurso de la cuarta clase y la entrega final se hará durante la última clase. En escala de 1 a 10, el trabajo se aprobará con una calificación igual o mayor a 5.

Informes
contacto@ie.org.ar
(5411)4382-2332
www.ie.org.ar

