

Introducción a R para Ciencias Sociales

Estación Lastarria – 2023

Contexto general

En el mundo laboral actual, las instituciones públicas y privadas necesitan disponer de profesionales capacitados en técnicas computacionales, métodos estadísticos y teoría social, con el propósito de no solo enfrentar la tarea de analizar el conjunto de datos que se producen diariamente, sino que también, realizar modelamientos que posibiliten o permitan explicar y predecir una variedad de hechos sociales que son resultado del propio interés institucional.

Tomando en consideración lo descrito en el cuadro precedente, en el proceso de toma de decisiones institucionales, los equipos de trabajo profesionales deben estar preparados tanto en el manejo de softwares computacionales (con la finalidad de abordar diversas aplicaciones de métodos funcionales relacionados a problemas específicos de las actividades laborales) así como en el fortalecimiento de capacidades que contribuyan a desarrollar procesos de análisis de datos y promover el razonamiento lógico.

En efecto, la entrega de reportes periódicos, la planificación estratégica de acciones, el uso de herramientas estadísticas como R, la confección de indicadores de gestión, el asesoramiento estadístico, la sistematización de los datos y el procesamiento de información fidedigna, son aspectos que cruzan todas las actividades institucionales. Y, frente a esto, los profesionales de las Ciencias Sociales pueden realizar un aporte importante.

Descripción

El presente programa tiene por objetivo entrenar en metodologías y competencias computacionales, el manejo de volúmenes de información y la subsecuente problematización a partir de las ciencias sociales. Todo ello con el propósito de apoyar en la gestión institucional de las labores profesionales que tienen los alumnos/as en sus campos de trabajo y/o estudio.

Responsable

Sebastián Massa Slimming

- Cientista Político. Magíster en Ciencias Sociales mención Sociología de la Modernización, Universidad de Chile. Diplomado en Estadística, Pontificia Universidad Católica de Chile, Diplomado en Estudios Latinoamericanos, Universidad de Santiago. Certificación en R Programming por John Hopkins University y Data Wrangling por Harvard University, además de poseer Licencia de Statistics Fundamentals with R. Acreditación en Licencia de Probabilidad y Estadística y Programación Estadística por la Universidad Nacional Autónoma de México. Actualmente ejerce labores de docencia a nivel de postgrado en Facultad de Ciencias Sociales, Universidad de Chile y a nivel de Educación Continua por la Facultad de Matemáticas, Pontificia Universidad Católica de Chile. A su vez, se desempeña como analista de datos por la Facultad de Matemáticas y DATA UC de la misma casa de estudios.

Público objetivo

Estudiantes de pregrado (cuarto año en adelante) y postgrado del área de las Ciencias Sociales, Humanidades, y Ciencias de la Empresa, Marketing y Administración. Profesionales de las mismas áreas mencionadas.

Requisitos mínimos

- Se requiere un conocimiento general de metodología y técnicas de investigación.
- Cada estudiante deberá contar con un computador portátil operativo.
- No se requieren conocimientos computacionales previos excepto un manejo intermedio en instalación y configuración de softwares (descarga de archivos, instalación y configuración de programas, etc.).
- Es deseable un manejo intermedio del idioma inglés.

Objetivos

Objetivo general:

Conocer las generalidades del software estadístico R y sus principales aplicaciones en el ámbito de las Ciencias Sociales, con la finalidad de realizar operaciones de análisis de datos en el contexto de la era del Big Data.

Objetivos específicos:

1. Identificar las principales características del lenguaje de programación en R.
2. Realizar cálculo de estadísticos descriptivos y representaciones gráficas.
3. Capacitar en la investigación y aplicación de diversas herramientas de Data Science.
4. Identificar los principales conceptos de estadística descriptiva.
5. Ejecutar operaciones para manipular y operar bases de datos.

Metodología

- Clases teóricas/expositivas.
- Talleres de metodologías para entrenamiento en competencias computacionales y razonamiento lógico.
- Desarrollo de líneas de código en software especializado.
- Análisis de casos y discusión teórica.
- Talleres y laboratorios de aplicación mediante uso de bases de datos contextualizadas.

Desarrollo de talleres

Se realizarán 1 actividad individual y 2 talleres grupales. Los talleres consistirán en ejercicios prácticos basados en análisis de bases de datos, abordando la teoría y metodología de programación en lenguaje R. Las evaluaciones grupales serán guiadas por el académico responsable durante la misma sesión de clases.

Contenidos por sesión

Sesión¹	Contenidos	Fecha
1. Introducción a la programación en R	1.1. Descripción de software R. 1.2. Lenguaje orientado a objetos. 1.3. Interfaz (prompt, editor y menú). 1.4. Tipo de variables. 1.5. Operadores aritméticos. 1.6. Operadores lógicos. 1.7. Operadores relacionales. 1.8. Estructura de datos cualitativos y cuantitativos. 1.9. Actividad individual	11-01-2023
2. Operación de base de datos en R.	2.1. Introducción a package Tidyverse. 2.2. Crear, probar y grabar bases de datos. 2.3. Concatenar y separar archivos (registros y variables). 2.4. Transformar y recodificar variables.	12-01-2023
3. Operación de base de datos en R. Taller Práctico I	3.1. Cruzar base de datos. 3.2. Manipulación de base de datos. 3.3. Limpiar base de datos. 3.4. Taller práctico grupal	13-01-2023
4. Introducción a estadística descriptiva.	4.1. Medidas de tendencia central. 4.2. Medidas de dispersión. 4.3. Estadísticos de forma. 4.4. Cálculo de estadísticos de resumen.	14-01-2023
5. Taller Práctico II	5.1. Taller práctico grupal	18-01-2023
6. Análisis, funciones y loops en R.	6.1. Estructuras de control (if, else, while, for, break, next). 6.2. Familia apply. 6.3. Introducción a visualización de datos.	19-01-2023

Bibliografía sugerida

- Healy, K. Data Visualization: A Practical Introduction." . Princeton University Press. 1a edición. (2019).
- James, G. An Introduction to Statistical Learning, with Applications in R". Springer. (2013).
- Saltz, J. S & Stanton, J. M. An introduction to Data Science." . Los Angeles: SAGE. (2018).
- Ware, C. Information Visualization: Perception for Design. O'Reilly". Waltham, Ma: Morgan Kaufman.. 3ra edición. (2013).
- Wickham, H. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data" Sebastopol, CA: O'Reilly Media. ISBN 1491910399. (2017).
- Wickham, H. R Packages. Sebastopol, CA: O'Reilly Media. ISBN 978-1491910597. (2017).
- Wilke, C. Fundamentals of Data Visualization. O'Reilly". O'Reilly Media. 1a edición. (2019).
- Xie, Y., Allaire, J.J. y Golemund, G. R Markdown: The Definitive Guide". CRC Press. (2019).

¹ La duración de cada sesión son 2,5 horas cronológicas.