

ESCUELA DE VERANO

TALLER DE INTRODUCCIÓN AL PROCESAMIENTO Y ANÁLISIS DE DATOS SOBRE MERCADO DE TRABAJO CON R

Coordinación

Dr. Francisco Nicolás Favieri (CONICET - GEIS - IISE - FACSO - UNSJ)

Dra. Beatriz Soria (FCPyS - UNCuyo)

Mail de consulta: cscomputacional@gmail.com

Del 24 al 28 de febrero de 15 a 18 hs - Modalidad virtual

FUNDAMENTACIÓN

El avance de herramientas digitales para el procesamiento y análisis de datos sobre mercado de trabajo y conflictividad laboral ha generado nuevas oportunidades y desafíos en la investigación social. Sin embargo, su complejidad y amplitud requieren la consolidación de una comunidad de estudios dedicada al uso de estas tecnologías, particularmente el lenguaje de programación R y su entorno para el análisis estadístico, la visualización de datos y la ciencia de datos.

Este curso está dirigido a personas con conocimientos básicos en estrategias cuantitativas aplicadas al estudio del mercado de trabajo, empleo y conflictividad laboral. R, como lenguaje de programación libre, permite desde análisis descriptivos hasta modelado avanzado, ofreciendo un ecosistema de código abierto respaldado por una comunidad global activa y recursos actualizados.

Además de su capacidad analítica, R destaca por la creación de visualizaciones que facilitan la comunicación de resultados complejos. Su uso es esencial en la investigación académica, la formulación de políticas públicas y el mercado laboral actual.

La enseñanza de R en este curso se sitúa en un contexto interdisciplinario que combina la ciencia de datos con las ciencias sociales computacionales. Este enfoque permite analizar fenómenos sociales a gran escala, realizar inferencias predictivas y expandir las fronteras de la investigación social mediante el uso de métodos híbridos.

El dominio de R no solo proporciona herramientas técnicas para el análisis de datos en el estudio del mercado de trabajo, sino que también responde a una demanda creciente en diversos sectores, fortaleciendo las capacidades para abordar problemas complejos con una perspectiva basada en datos.

OBJETIVOS

Objetivo General

- Introducir a lxs participantes en el uso del lenguaje de programación R para el procesamiento y análisis de datos aplicados a indicadores socioeconómicos, con énfasis en el mercado de trabajo.

Objetivos Específicos:

- Comprender los procesos de relevamiento, análisis y producción de información en estudios socioeconómicos.
- Explorar y manipular bases de datos pertinentes. Encuesta Permanente de Hogares (EPH).
- Reflexionar críticamente sobre la producción de información estadística y sus usos metodológicos.
- Consolidar “buenas prácticas” en la manipulación, transformación y visualización de datos como en la elaboración y difusión de resultados de investigación.
- Fomentar el trabajo colaborativo mediante el uso de herramientas computacionales para el análisis de datos en equipos de investigación.

DESTINATARIXS

Esta propuesta está destinada a integrantes de equipos de investigación vinculados a los estudios del trabajo y las ciencias sociales computacionales que participen en los proyectos de investigación y extensión, grupos de estudio y cátedras seleccionadas en el marco del Núcleo de Ciencias Sociales Computacionales (Res. en trámite). Se libera la inscripción a 10 personas externas.

CONTENIDOS

Encuentro 1: Introducción a R y las Ciencias Sociales Computacionales

Introducción a las Ciencias Sociales Computacionales y a la programación con R. Definiciones. Bases de datos para estudios del trabajo: EPH. Programas: Conceptos básicos. Instalación R. Origen. IDEs. IDEs: pantallas de trabajo, estructura de funcionamiento y paquetes. Ejercicios.

Encuentro 2: RBase yRStudio

Operaciones básicas en R: apertura y cierre, gestión de carpetas de trabajo, gestión y tipo de archivos y comandos básicos. Elementos: objetos, vectores, tablas, bases de datos. Características de variables. Revisión de paquetes básicos. Ejercicios.

Encuentro 3: Manipulación, transformación y procesamiento de datos

Presentación de paquetes (Tidyverse, eph). Manipulación de bases de datos: control de casos y estructura (medidas resumen). Ordenamiento, definición y creación de variables. Unificación y selección de datos. Procesamiento de datos: de datos: valores faltantes, recodificación y renombramiento de variables. Selección, ordenamiento y agrupación de datos. Ejercicios.

Encuentro 4: Visualización de datos en R

Visualización de datos. Representación de datos en tablas y gráficos. Creación de tablas. Gráficos según tipo de datos, edición y ejecución. Revisión de paquetes (ggplot2, kable). Ejercicios.

Encuentro 5: Producción de informes y automatización de reportes

Introducción a R Markdown para la creación de informes reproducibles. Creación de documentos HTML, PDF y Word desde RStudio. Integración de gráficos y tablas en reportes dinámicos. Exportación y automatización de resultados de investigación.

METODOLOGÍA

El taller se desarrollará en cinco encuentros virtuales de tres horas cada uno, con dos pausas de 15 minutos.

El enfoque será teórico-práctico, combinando la exposición de conceptos con la implementación en R. Cada sesión incluirá:

- Explicación teórica sobre conceptos clave y herramientas de R.
- Ejercicios prácticos guiados en vivo, aplicados a datos reales del mercado de trabajo.
- Discusión y análisis de resultados con enfoque crítico.

Para favorecer el aprendizaje, se trabajará con bases de datos socioeconómicas, como la Encuesta Permanente de Hogares (EPH). Se incentivará la participación activa y el trabajo colaborativo mediante actividades grupales y resolución de casos aplicados.

Criterios de aprobación

Para obtener la certificación, se requiere:

- Asistencia mínima del 75% (al menos cuatro encuentros).
- Presentación y aprobación de un trabajo final integrador en grupo, que consistirá en el análisis de una base de datos utilizando las herramientas aprendidas.
- Se proporcionará material de apoyo a través de un repositorio en GitHub, incluyendo código, lecturas y ejercicios complementarios.

REFERENCIAS BIBLIOGRÁFICAS Y FUENTES DOCUMENTALES

Materiales del curso

- Repositorio en **GitHub** con materiales, códigos y ejercicios (se proporcionará el enlace cuando esté disponible).

Fuentes oficiales de datos

- Instituto Nacional de Estadística y Censos de la República Argentina (INDEC). (2019). *Encuesta Permanente de Hogares: Nota metodológica. Primer trimestre de 2019*. [Disponible aquí](#).

Bibliografía general

- Soria, B., Abarzúa Cutroni, A., Favieri, F. N., y Rosati, G. F. (2024). El riesgo está del lado de la crítica abstracta más que del retorno al empirismo ingenuo. *Tramas Sociales: Revista del Gabinete de Estudios e Investigación en Sociología (GEIS)*, 6(6), 98–113. Recuperado de <https://ojs.unsj.edu.ar/index.php/tramassociales/article/view/1260>
- Vázquez Brust, A. (2021). *Ciencia de Datos para gente sociable* (Cap. 1, 2). [Disponible aquí](#).
- Wickham, H. (2017). *R para Ciencia de Datos*. O'Reilly. (Cap. 1, 2, 6, 17, 20). [Disponible aquí](#).
- Wickham, H. (2019). *The Tidyverse Style Guide* (Cap. 4). [Disponible aquí](#).
- Xie, Y., Dervieux, C., & Riederer, E. (2020). *R Markdown Cookbook*. Chapman & Hall/CRC. [Disponible aquí](#).