

Introducción al Programa:



Duración: 24 horas (presencial).



Inicio: 05/agosto/2024, **termina:** 14/agosto/2024. Ocho sesiones, 9-12 horas (GMT-7).

Cupo: para un mínimo de 1 y máximo de **20 personas**.



Lugar: El Colegio de Sonora, Plaza Hidalgo en Calle Obregón 54, colonia Centro, Hermosillo, Sonora, México. Sala 3 del Edificio de Postgrado, entrada por Callejón Ayón.



Créditos como curso de **“Temas selectos”** o se expide certificado.



Costo: Gratuito a estudiantes de El Colegio de Sonora.



Instructor: Dr. Luis Alan Navarro Navarro

Introducción:

Mi primer contacto con el programa R fue alrededor de hace 18 años, durante mis estudios doctorales, cuando, aunque para ese entonces ya había llevado al menos cuatro cursos de estadística donde había usado cuatro diferentes programas/software, me encontraba en la disyuntiva de ¿qué programa usar para el manejo y análisis de datos de mi disertación? Un compañero, estudiante de doctorado y especialista en sistemas computacionales me aconsejó: usa R, es gratuito, está evolucionando rápidamente, cada vez más investigadores y académicos lo están adoptando.

Ese fue uno de los mejores consejos que alguien me ha dado. Desde entonces he trabajado con R en la más amplia diversidad de proyectos; sin desestimar las capacidades de otros programas de análisis de datos (SAS, SPSS, STATA, etc.), nunca he tenido la necesidad de usarlos; R ha resuelto satisfactoriamente mis necesidades.

Y a todo esto ¿Qué es R? Es un lenguaje de programación orientada a objetos para el análisis estadístico, matemático y la creación de gráficos. Fue creado por **Ross Ihaka** y **Robert Gentleman** en 1991 en la Universidad de Auckland en Nueva Zelanda (de ahí el nombre de “R”). Es un

software abierto que se puede descargar gratuitamente de Internet. La primera versión de R se lanzó el año 2000, evolucionando muy rápido, **actualmente (2024) va en la versión 4.3.1.**

R está en constante evolución. Existe una amplia y diversa comunidad científica y de desarrolladores de software que, a como lee estas líneas, están creando nuevas librerías que aumentan la capacidad y poder analítico del programa. Por ejemplo, al inicio R tenía alrededor de 100 librerías, actualmente (junio de 2024) [hay 20,923 librerías disponibles en el repositorio CRAN \(Comprehensive R Archive Network\)](#). En su conjunto, estos programas permiten trabajar con una gran diversidad de bases de datos.

Adicionalmente, existen una cantidad enorme de tutoriales, blogs, videos y comunidades que comparten soluciones técnicas a problemas y errores que surgen en el uso de los paquetes. En 18 años de ser usuario asiduo, siempre he encontrado la solución a un problema o reto analítico en Internet. [Hoy en día con la Inteligencia Artificial tenemos prácticamente a un asesor con nosotros siempre.](#)

Aprender R es un **proceso continuo**, no un destino o meta que se alcanza. Constantemente emergen nuevas librerías con herramientas más novedosas y eficientes. Hoy en día, como sociedad estamos generando datos a raudales que requieren de ser analizados y simplificados, invertir tiempo en aprender R, es algo que, como yo, no va a arrepentirse.

Para quien esta diseñado este curso:

Este curso esta pensado para estudiantes de nivel superior (licenciatura) y de postgrado (maestría o doctorado) que tienen la necesidad de analizar diversa información cualitativa y/o cuantitativa para el desarrollo de sus trabajos de tesis o disertación.

Muy frecuentemente, los asesores y/o directores de investigación utilizan R, por lo que exigen que sus estudiantes posean un nivel básico de este programa, para desarrollar después análisis más especializados.

Requisitos:

La experiencia previa con conceptos y ejercicios de estadística descriptiva es deseable y ayuda bastante. Pero para este curso, no es necesario tener conocimientos previos de estadística o el manejo de otros programas computacionales para el análisis de datos. Se requiere que el alumno sea capaz de navegar por Internet, especialmente en idioma inglés y que pueda comprender texto en inglés especializado (técnico). Es necesario tener conocimientos básicos en el uso de ordenadores (computadoras con plataforma Windows o Mac). La curva de aprendizaje inicial de R es muy demandante, principalmente requiere de mucha paciencia, insistencia y dedicación inicial. Muchos de los errores cometidos al inicio son por erratas o errores tipográficos al momento de estar escribiendo las instrucciones y comandos de R, el editor de RStudio es de gran ayuda, pero aún así no es infalible. El estar familiarizado con el uso de herramientas de Inteligencia Artificial (IA) es de gran ayuda, sin embargo se recomienda que primero el estudiante conozca R para después poder evaluar las salidas generadas por IA como ChatGPT.

Programa del curso:

1 Introducción (00:40)

- 1.1 Bienvenida y (auto) presentación del profesor.
- 1.2 Presentación: [¿Por qué usar R y no otro \(Software\)?](#)

2 RStudio (00:30)

- 2.1 Presentación: [¿Qué es RStudio?](#)
- 2.2 Alternativamente se puede usar RStudio Cloud.

3 Crear un proyecto en RStudio (00:30)

- 3.1 Video tutorial: [Crear un proyecto en RStudio.](#)

4 Creamos nuestro primer código (00:40)

- 4.1 Creamos un Snippet para el encabezado del archivo para código.
- 4.2 [Descarga este archivo de extensión txt.](#)
- 4.3 Presentación: [Crear un archivo para código. Uso de Snippets.](#)
- 4.4 Video: [Crear un archivo para código. Uso de Snippets.](#)

5 Tipos de objetos, creación, indexación, acceso (4:00)

- 5.1 ¿Por qué objetos?
- 5.2 ¿Cómo nombrar objetos?
- 5.3 Objetos en el espacio del proyecto.
- 5.4 Objetos para datos.
- 5.5 Tipos de variables.
 - 5.5.1 Numéricas.
 - 5.5.2 Alfanuméricas.
 - 5.5.3 Valores perdidos.
 - 5.5.4 Factores.
 - 5.5.5 Fechas.
- 5.6 Atributos de los objetos; acceso, indexación, operaciones aritméticas.
 - 5.6.1 Vectores.
 - 5.6.2 Matrices.
 - 5.6.3 Variables indexadas ("array").
 - 5.6.4 Lista.
- 5.7 Coerción de tipos de variables.
- 5.8 Limpieza de datos alfanuméricos.
- 5.9 Variables y operadores lógicos.

6 Cómo instalar/cargar paquetes en R (0:30)

- 6.1 Tutorial: [Instalar y cargar paquetes en R.](#)
- 6.2 Video: [Instalar y cargar paquetes en R.](#)

7 Gestión de bases de datos (4:00)

- 7.1 Creación de un *data.frame*.
- 7.2 Importar hojas de datos de diferentes formatos.
- 7.3 Crear Snippet para cargar un "dataframe"
 - 7.3.1 Creamos un Snippet leer un archivo "csv".
 - 7.3.2 [Descarga este archivo de extensión txt.](#)

- 7.3.3 [Liga a la base de datos "tomas"](#).
- 7.3.4 Video: [Crear un Snippet para cargar un "dataframe"](#).
- 7.3.5 Video: [Cómo importar un dataframe con botones](#).
- 7.4 Atributos de las hojas de datos.
- 7.5 Manejo de hoja de datos.
 - 7.5.1 Acceso a variables.
 - 7.5.2 Tipo de variables.
 - 7.5.3 Creación de nuevas variables.
 - 7.5.4 Recodificación.
 - 7.5.5 Agregar registros (hileras).
- 7.6 Reestructuración de datos.

8 Manipulación de hojas de datos con *Tidyverse* y *Dplyr* (4:00).

- 8.1 La función `filter()`
- 8.2 La función `select()`
- 8.3 La función `mutate()`
- 8.4 La función `arrange()`
- 8.5 La función `summarize()`
- 8.6 Agrupar y resumir datos.
- 8.7 La función `group_by()`
- 8.8 Resúmenes agrupados con `summarize()`
- 8.9 Combinando `group_by()` con otras funciones `dplyr`
- 8.10 Unir bases de datos.
- 8.11 Las funciones `left_join()`, `right_join()`, `inside_join()` y `full_join()`
- 8.12 Uso avanzado del operador de tuberías `%>%`
- 8.13 Combinando múltiples funciones `dplyr` en una tubería

9 “Reproducibilidad” del análisis (0:30)

- 9.1 Video: [Test: “Reproducibilidad” del análisis](#).

10 Elaboración de gráficas y análisis descriptivo básico (2:40)

- 10.1 Estadística descriptiva básica.
- 10.2 Estadística descriptiva básica.
- 10.3 Tablas (variables categóricas).
- 10.4 Dos formas de graficar: `plot()` y `ggplot2`.
- 10.5 Graficas elementales.
- 10.6 Personalización de gráficas.
- 10.7 Exportar gráficas.

11 Breve introducción a R Markdown (4:00)

- 11.1 ¿Qué es R Markdown y cuál es su utilidad?
- 11.2 Lenguaje y estructura de R Markdown.
- 11.3 *Chunks* o fragmentos de código R.
- 11.4 R Notebook.
- 11.5 Crear un R Notebook.

Sobre el profesor del curso:

Dr. Luis Alan Navarro Navarro

Desde el año 2014 es profesor investigador en El Colegio de Sonora en el Centro de Estudios en Gobierno y Asuntos Públicos, ubicado en la ciudad de Hermosillo, Sonora; adscrito desde el año 2014 al Programa Investigadoras e Investigadores por México CONAHCYT, desarrollo dos líneas de investigación: a) gestión integral del agua en zonas áridas; y b) gestión sustentable del bosque urbano. Coordinador de la [Maestría en Gestión Integral del Agua](#).

Es ingeniero agrónomo zootecnista por la Universidad de Sonora; maestro en economía del desarrollo rural por la Universidad Autónoma Chapingo; y doctorado en Ciencias de los Recursos de Zonas Áridas (Programa Multidisciplinario) por la Universidad de Arizona el año 2012. Desde mis estudios doctorales he utilizado el programa R en mi disertación doctoral y después en todos mis trabajos de investigación y proyectos de consultoría. No puedo esperar el siguiente proyecto en el que tenga que volver a usarlo, disfruto el reto que impone el crear y modelar los datos en este programa. Puedes visitar mi página personal alanphd.com en Internet para saber más sobre mi trabajo.