



CURSO INTRODUCCIÓN AL SOFTWARE ESTADÍSTICO R

División: Ciencias Económicas y Administrativas / Modalidad: Presencial

**INSCRIPCIONES
ABIERTAS**

Propósito de formación:

El curso proporciona al estudiante las competencias para el manejo adecuado del software R aplicadas al tratamiento y organización de datos.

FECHA DE INICIO: 11 de junio de 2019
FECHA DE FINALIZACIÓN: 11 de julio de 2019
HORARIOS: martes y jueves 6:00 p.m. a 10:00 p.m.

DURACIÓN: 40 horas
INVERSIÓN: \$800.000



MODULOS

INTRODUCCIÓN (10 horas)

- Historia de R, descripción general del programa (ventajas y desventajas)
- Scripts y Directorios de trabajo
- Cómo cargar los datos en R en distintos formatos.
- Tipos de datos
- Clases de objetos (Vectores, Matrices, Listas y data frames)
- Operadores básicos en la programación en R (Operadores algebraicos, lógicos, relacionales, etc)
- Creación de funciones en R y estructuras de control: loops, if, else.

MANEJO DE INFORMACIÓN ESTADÍSTICA: (15 horas)

- Bases de datos, estructuras de datos en R, importación y exportación de datos, agregación de datos, unión de tablas (merging), recodificación.
- Consultas y reestructuración de datos: consultas a través de las funciones básicas de R, uso del paquete RSQLite.
- Reestructuración de bases de datos (reshaping).
- Manejo del paquete de R dplyr.
- Manejo de diferentes formatos de datos: operadores lógicos, manejo de subíndices, manejo de datos tipo factor, fecha y caracteres. Manejo del paquete tidyr.

HERRAMIENTAS BÁSICAS DE VISUALIZACIÓN Y ANÁLISIS EXPLORATORIO DE LOS DATOS (15 horas)

- Visualización de variables categóricas: proporciones, tablas de contingencia (perfiles fila y columna) en R
- Visualización gráfica de variables continuas y cruces de variables: histogramas, diagrama de tallos y hojas, diagrama de caja, técnicas de detección de atípicos, diagramas de dispersión, matriz de dispersión, uso de diagrama de cajas en conjugación con variables categorías.
- Gramática de gráficos: uso de los paquetes ggplot2 y lattice.

