

UNIVERSIDAD NACIONAL DEL COMAHUE

Facultad de Economía y Administración

Maestría en Estadística Aplicada



***Curso de Posgrado
Aprendiendo R con Estadística***

Docentes a cargo: Dr. Gustavo Giménez - Mg. Natalia Rubio

Docente colaboradora: Lic. Ana Haique

Neuquén

2023

Denominación: Aprendiendo R con Estadística

TIPO DE CURSO: Curso de posgrado

REQUISITOS DE ADMISIÓN: Título de grado de al menos cuatro años de duración. Conocer los contenidos básicos de un curso de Estadística de grado (estadística descriptiva, probabilidad, modelos de probabilidad, inferencia estadística, análisis de regresión lineal, correlación)

JUSTIFICACIÓN DEL CURSO

R es un ambiente de programación libre de distribución gratuita que, permite implementar una vasta gama de análisis estadísticos clásicos, gráficos de alta calidad y también técnicas modernas como machine learning o statistical learning. Actualmente, es uno de los softwares más utilizados para análisis estadísticos, en investigación y ambientes académicos en general. Es una herramienta de trabajo en variadas disciplinas como econometría, agronomía, geografía, entre otros. El objetivo general del curso es: brindar las herramientas básicas del software R y, a partir de los conceptos de estadística de los cursos de grado, explorar las diferentes funciones y librerías.

OBJETIVOS:

Los objetivos específicos del curso son los siguientes:

- Transmitir a los estudiantes los conceptos básicos del lenguaje R.
- Entrenar a los alumnos en la aplicación de herramientas tanto de la estadística descriptiva como inferencial, con problemas típicos de análisis de datos en economía y en estudios agronómicos, utilizando R.
- Brindar elementos de programación en R para la creación básica de programas y de funciones propias.
- Lograr en los estudiantes la autonomía necesaria que permita identificar los alcances y nuevas posibilidades que se abren dentro del lenguaje R.

CARGA HORARIA: 60 horas

PROGRAMA SINTÉTICO:

Introducción al ambiente R. Estadística descriptiva, análisis univariado y Manejo básico de R. Creación de gráficos utilizando funciones base. Análisis descriptivo bivariado y Multivariado. Inferencia. ANOVA. Regresión y correlación.

PROGRAMA ANALITICO:

1. Introducción al ambiente R. Que es R y por qué utilizarlo con Estadística? Instalación. Inicio y cierre de sesión. Espacio de trabajo. Librerías. Repositorios. Descargas. Espejos. Interfaz Rstudio. Proyectos. Ayudas. ChatGPT. Manuales. Foros. Sitios de interés.
2. Estadística descriptiva, análisis univariado y Manejo básico de R. Tipo de variables: objetos en R, Vectores enteros, numéricos, y de caracteres, factores. Selección de Vectores. Indicador de

datos perdidos. Creación y generación de vectores. Medidas resumen. Funciones definición y uso, argumentos. Elementos básicos de programación: Funciones propias.

3. Creación de gráficos utilizando funciones base: Gráficos: Histogramas. Gráficos de densidad. Gráficos de puntos. Gráficos para variables enteras. Gráficos de sectores. Q-Q Plot. Gráficos de cajas. Grafico de barras. Mosaicos.
4. Análisis descriptivo bivariado y Multivariado. Arreglos y matrices, operaciones. Archivos de datos. Tablas. Listas. Importar y exportar datos. Gráficos bi y multivariados. Introducción a tidyverse y gráficos en ggplot2.
5. Inferencia: Prueba de hipótesis t para una y dos muestras independientes. Prueba t para muestras apareadas. Pruebas de independencia chi-cuadrado. Pruebas sobre Varianzas poblacionales.
6. ANOVA: Análisis de la varianza de un factor, comparaciones múltiples. Errores. DBCA. Anova con estructura factorial de tratamientos.
7. Regresión y correlación: Correlación. Regresión lineal simple. Residuales y valores ajustados. Intervalos de confianza y predicción. Series de tiempo. Transformaciones. Descomposición de las series.

BIBLIOGRAFIA BASICA:

Baumer, B. S., Kaplan, D. T., Horton N. J. (2021) Modern Data Science with R. CRC press.

Crawley, M.J. (2007). The R Book. Wiley.

Cowpertwait, P., Metcalfe, A. (2009). Introductory Time Series with R. Springer.

Dalgaard, P. (2008). Introductory statistics with R. 2nd ed. Springer.

Faraway, J.J. (2009). Linear models with R. Chapman & Hall/CRC Press, Florida, USA.

Gujariti, D., Porter, D. (2005). Econometría. Quinta Edición. McGraw Hill.

Coghlan, A. (2018). A Little Book of R For Time Series. Trust Sanger Institute, Cambridge, U.K

Lander, P.J. (2017) R for everyone. Advanced Analytics and Graphics. Second Edition.

Paradis, E. (2005). R para principiantes. cran.r-project.org/doc/contrib/rdebuts.es.pdf

Venables, W.N.; Smith, D.M. (2014). An introduction to R. Notes on R: A programming Environment for data analysis and graphics. cran.r-project.org/doc/manuals/R-intro.

Wickham, H. (2020). ggplot2: Elegant Graphics for Data Analysis. Springer.

Zuur, A.F., Ieno, E.N., Smith, G.M. (2007). Analyzing ecological data. Springer, New York.

Otras referencias:

Cendoya, M.G.; Monterubbianessi, G.; Giménez, G. (2016); Introducción a R. Apuntes del curso.

Romero, M., Giménez, G. (2021); Principios de Programación en R. Versión Previa. <http://dx.doi.org/10.13140/RG.2.2.10272.25601>

[doi.org/10.13140/RG.2.2.10272.25601](http://dx.doi.org/10.13140/RG.2.2.10272.25601)

Material adicional será provisto durante el curso.

EVALUACIÓN:

Asistir al menos al 80% de las clases.

Aprobar dos trabajos prácticos de resolución de problemas utilizando las técnicas abordadas.

Aprobar los cuestionarios mediante el uso de la PEDCO con una exigencia del 60%.

CERTIFICACIÓN QUE SE EXPIDE: Curso de Posgrado "Aprendiendo R con Estadística".

MODALIDAD:

Los encuentros serán presenciales. Serán de carácter teórico-prácticos con desarrollo de teoría y trabajo individual en computadoras utilizando R para resolver casos de estudio.

FECHA TENTATIVA:

El curso se dictará los viernes de 18 a 22 h y sábados de 8 a 12 h. La fecha tentativa de inicio es en Septiembre.

ARANCELES:

Arancel general: \$10.000.-

Arancel docentes departamento de Estadística: \$ 5.000.-

CUPO: 30 participantes