**Curso de Postgrado**

**Introducción a R para el análisis exploratorio de datos con métodos multivariados**

**9 al 13 de mayo de 2016**

**Organizado por el Instituto de Hidrología de Llanuras “*Dr. Eduardo J. Usunoff*”**

1. **Docente**

***Priscilla G. Minotti***

Co-Directora del Laboratorio de Ecología, Teledetección y Ecoinformática y Profesora Asociada del Instituto de Investigación e Ingeniería Ambiental, Universidad de San Martín, Gral. San Martín, Buenos Aires, Argentina.

MSc Geography, Oregon State University, (USA), programa de Técnicas Geográficas; Esp. Explotación de Datos y Gestión de Conocimiento, UBA; Lic. Cs. Biológicas, FCEyN UBA.

1. **Objetivos**

Este curso tiene como objetivo introducir al entorno R y aprender a usarlo para realizar análisis exploratorio de datos mediante métodos multivariados.

El curso parte de la descarga, instalación y configuración del entorno R y presenta terminología y conceptos básicos de R como lenguaje de programación para el análisis de datos y para la escritura de funciones. Cubre aspectos prácticos de carga de datos, acceso a paquetes, preprocesamiento de datos y salidas gráficas. Introduce distintos enfoques de análisis exploratorio de datos, los paquetes de R más comúnmente usados y ejemplos de aplicación.

1. **Programa Analítico**

***Día 1* (09/05/2016)**

**R** como entorno de programación y análisis de datos. Una breve historia. Descarga e instalación de **R**. Descarga e instalación de RStudio. Instalación de paquetes. La interfase de **R** y RStudio. Entrada y evaluación a través de la consola de R. El directorio de trabajo. Cómo pedir ayuda sobre un tema. Objetos de **R**. Almacenamiento y eliminación de objetos. Atributos (nombre, modo, clase, dimensiones). Funciones. Librerías. Carga de paquetes. La interfase de RStudio.

Datos en **R**. Datos básicos o atómicos. Vectores. Listas. Matrices. Factores. Valores faltantes. Data frames (hojas de datos). Nombres de atributos.

Gestión básica de datos. Lectura e importación de datos. Generación de secuencias de datos. Creación y modificación. Índices. Operaciones simples. Concatenación. Selección y eliminación de datos con valores particulares. Funciones para datos agrupados por factores.

Procedimientos gráficos. Análisis estadísticos básicos.

***Día 2* (10/05/2016)**

Análisis exploratorio de datos multivariados. Tipos de datos. Descripción estadística de datos multivariados. Datos atípicos. Medidas de asociación y distancias. Datos faltantes. Variables redundantes. Transformaciones lineales. Transformaciones no lineales.

Distribuciones multivariadas. Métodos factoriales, de “ordenamiento” o reducción de información: Componentes principales, Escalado multidimensional, Análisis de correspondencias. Métodos de agrupamiento o análisis de conglomerados: Ordenamiento jerárquico, Ordenamiento no jerárquico.

Prácticas de lectura, organización y descripción estadística de datos multivariados en **R**.

***Día 3* (11/05/2016)**

Análisis exploratorio de datos multivariados en entorno **R**.

El concepto de resultados reproducibles en investigación científica. Planteo de objetivos. Datos disponibles. Selección de métodos apropiados a la pregunta y los datos. Organización del flujo de datos e información. Almacenamiento de la secuencia de funciones utilizadas como objeto de **R**. Preparación de datos. Análisis. Salidas numéricas y gráficas. Interpretación.

Prácticas con datos de distintos casos de estudio.

***Día 4* (12/05/2016)**

Trabajo con datos propios. Planteo de objetivos. Datos disponibles. Selección de métodos apropiados a la pregunta y los datos. Paquetes disponibles. Organización del flujo de datos e información. Preparación de datos propios. Descripción estadística. Análisis exploratorio multivariado. Salidas numéricas y gráficas. Interpretación. Secuencia de funciones **R** utilizadas.

***Día 5* (13/05/2016)**

Presentación de trabajos individuales y/o grupales.

1. **Carga horaria**

40 horas (5 días de 8:30 – 12:30, y de 14:00 – 18:00)

1. **Requerimientos.**

Cada participante deberá traer su computadora portátil a fin de realizar la instalación de los programas a usar en el curso y también los paquetes de **R** que sean necesarios.

1. **Modalidad y Evaluación**

El curso es teórico-práctico con modalidad presencial y evaluación final dada por un proyecto de análisis exploratorio de datos realizado completamente por el/los alumnos.

1. **Bibliografía**

# Borcard, Daniel; Gillet, Francois; Legendre, Pierre. 2011. Numerical Ecology with R. Springer.

Crawley, Michael J. 2012. The R Book, 2da. Ed. Wiley.

Langohr, Klaus 2010. Introducción a R. Universitat Politecnica de Catalunya. Disponible en <http://www-eio.upc.es/teaching/best/R/TutorialR.pdf>

Owen, W. J. 2006. The R Guide. <http://cran.r-project.org/doc/contrib/Owen-TheRGuide.pdf>

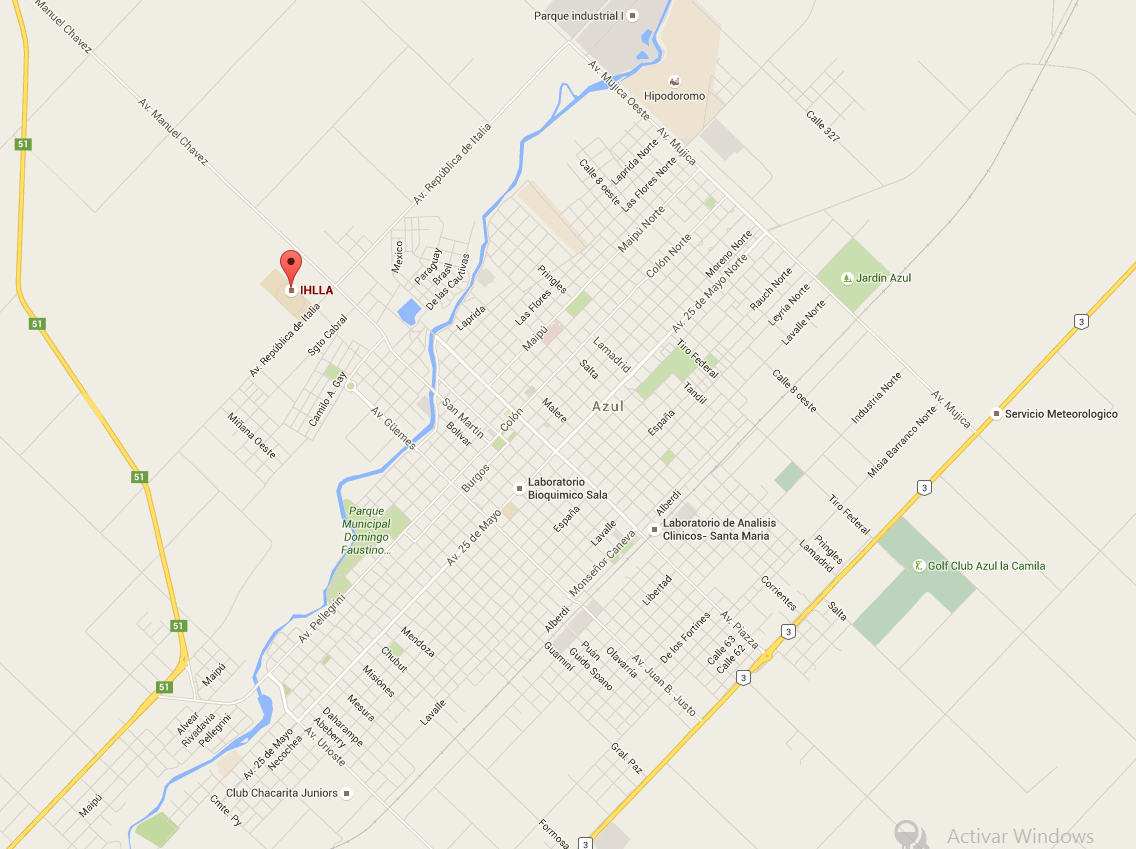
Paradis, Emmanuel. 2003. R para Principiantes. Disponible en <http://cran.r-project.org/doc/contrib/rdebuts_es.pdf>

R Development Core Team. 2000. Introducción a R. Disponible en <https://cran.r-project.org/doc/contrib/R-intro-1.1.0-espanol.1.pdf>

Short, Tom. 2004. R Reference Card. Disponible en <http://cran.r-project.org/doc/contrib/Short-refcard.pdf>

1. **Lugar**

Aula de postgrado del Instituto de Hidrología de Llanuras “*Dr. Eduardo J. Usunoff*”. Campus de la Universidad Nacional del Centro de la Provincia de Buenos Aires. Azul, Buenos Aires, Argentina.



1. **Consultas e información del curso**

Dra. ***Guadalupe Ares***

e-mail: [gares@faa.unicen.edu.ar](mailto:gares@faa.unicen.edu.ar)

Tel. +54 2281 432666 int. 43

1. **Inscripción y costo de matrícula**

Ing. (Ms.) ***Georgina Cazenave***

e-mail: [cazenave@faa.unicen.edu.ar](mailto:cazenave@faa.unicen.edu.ar)

Tel. +54 2281 432666 int. 35

**Costo de la matrícula**:

General $ 2.000 (dos mil); estudiantes de postgrado $ 1.500 (mil quinientos pesos).

El pago se realizará por transferencia bancaria a la cuenta:

Titular: Comisión de Investigaciones Científicas

CUIT: 33-68485414-9

Cuenta Fiscal Nº 1070/4 – Banco Provincia de Buenos Aires

Casa Matriz 2000 – La Plata – Argentina

CBU: 0140999801200000107043

Enviar comprobante y datos necesarios de facturación por fax (02281-432666) o e-mail: [cazenave@faa.unicen.edu.ar](mailto:cazenave@faa.unicen.edu.ar)

***Las plazas son limitadas, por lo que se recomienda inscribirse lo antes posible.***

***Información sobre la ciudad y hospedajes*:**

http://www.azul.gov.ar/

<http://www.conocelaprovincia.com.ar/azul-buenos-aires/>

<http://www.buenosairesturismo.com.ar/azul/alojamientos.php3>