

## Expe(R)iencias

### *temas en las cienCIAS de datos con R:*

#### RESUMEN

En la actualidad numerosas empresas y centros públicos utilizan R como herramienta de trabajo, lo que lo convierte en el líder del software libre con aplicación en Ciencias Sociales, Economía, Empresa y en la investigación biomédica.

El curso enseña a los alumnos a manejar R y a utilizarlo específicamente en la resolución de problemas habituales en el ámbito empresarial en las distintas áreas: financiera, administrativa, marketing y biomedicina.

El buen uso de R contribuirá a mejorar sus resultados tanto en el uso profesional como en las distintas asignaturas del grado (Matemáticas Financieras, Economía Financiera, Gestión Financiera, Estadísticas y Biomedicina). En particular, mejorará sin duda los informes y análisis económicos, financieros y de los ensayos clínicos de su carrera profesional.

Sin duda será una herramienta útil en la realización de vuestro TFG.

#### DIRIGIDO A

Los alumnos de primer curso de grado no pueden asistir debido a la falta de conocimientos básicos para poder seguir los contenidos del curso.

#### REQUISITOS

El curso no se impartirá en aulas informáticas por lo que se precisa traer ordenador.

**Los alumnos traerán su portátil con el siguiente software instalado y funcionando: R ([cran.r-project.org](http://cran.r-project.org)), R-Studio ([www.rstudio.com](http://www.rstudio.com)) y Keras R ([keras.rstudio.com](http://keras.rstudio.com) -- no es necesaria GPU). Durante el curso de darán instrucciones detalladas de la instalación de otros paquetes.**

Para recibir el certificado de asistencia se admitirá como mucho una falta de asistencia. Dicha falta será justificable. Dos o más faltas implicará por tanto no recibir certificado aunque el alumno podrá continuar asistiendo.

#### PROFESORADO

**J. Miguel Marín**

Profesor Titular del Departamento de Estadística, UC3M

**Stefano Cabras**

Profesor Ramón y Cajal del Departamento de Estadística, UC3M

# PROGRAMA

## SESION 1: Evolución del precio de la vivienda en Madrid.

- 1.1 Leer una tabla de datos de una fuente parcialmente organizada
- 1.2 Hacer gráficos informativos descriptivos
- 1.4 Modelización básica de una serie temporal
- 1.5 Aplicación a precios de la vivienda en Madrid

## SESION 3 Didáctica con R

- 3.1 Repaso de temas típicos de inferencia básica
- 3.2 Gráficos ilustrativos de conceptos de la inferencia
- 3.3 Análisis univariante y bivariante
- 3.4 El conjunto de paquetes statsTeachR

## SESION 5 Clasificación de cáncer con Google Tensor Flow

- 5.1 Instalar Google Tensor Flow
- 5.2 Redes neuronales como base del Deep Learning
- 5.3 Algoritmos de optimización
- 5.4 Aplicación a clasificación de cáncer

## SESION 2: Efecto de una campaña publicitaria

- 2.1 Introducción a la inferencia causal
- 2.2 El paquete Causal Impact
- 2.3 Modelos para efectos causales
- 2.4 Aplicación al impacto de una campaña causal

## SESION 4 Análisis de opiniones en twitter

- 4.1 Obtener los tweets
- 4.2 Filtrar los tweets con Tidy verse
- 4.3 Construir diccionarios de "sentimientos"
- 4.4 Representación de los contenidos por medio de Word Clouds

## SESION 6 Asociaciones genéticas en gran escala (estudios GWAS)

- 6.1 El concepto de contraste múltiple
- 6.2 Evidencia: p-valores y su calibración
- 6.3 Conceptos de False Discovery Rate y Local False Discovery Rate
- 6.4 Aplicación a una base de datos de cáncer de próstata

## FECHAS y HORARIO

Lunes 9 de abril, 10:00 – 13:00  
Miércoles 11 de abril, 10:00 – 13:00  
Viernes 13 de abril, 10:00 – 13:00

Lunes 16 de abril, 10:00 – 13:00  
Miércoles 18 de abril, 10:00 – 13:00  
Viernes 20 de abril, 10:00 – 13:00

## LUGAR DE LA CELEBRACIÓN

Universidad Carlos III de Madrid, Campus de Getafe.

## PRECIOS

Estudiantes y Ex – alumnos de la UC3M: 100 €

Estudiantes externos a la UC3M: 200 €

No Universitarios: 300 €

## FORMULARIO DE INSCRIPCIÓN

<https://goo.gl/DyV6nr>

## DIRECCIÓN DE CONTACTO

[etorijan@est-econ.uc3m.es](mailto:etorijan@est-econ.uc3m.es)