



## **Ciclo formativo en Big Data y Ciencia del Dato para la Estadística Oficial: ARQUITECTURA Y HERRAMIENTAS TECNOLÓGICAS PARA EL BIG DATA**

### **Profesorado:**

Antonio Beato Moreno  
Pedro Luis Luque Calvo

### **Objetivos:**

Introducir a los participantes en el estado de la tecnología para el procesamiento de grandes conjuntos de datos y aplicar los conocimientos adquiridos de forma práctica en máquinas virtuales mediante el uso de diferentes herramientas existentes en el ámbito del Big Data especializadas para tratar grandes conjuntos de datos.

### **Programa del curso:**

- 1) Visión general sobre las herramientas y tecnologías Big Data
- 2) Hadoop y MapReduce
- 3) Lenguajes para procesar y analizar datos en Hadoop: Pig y Hive
- 4) Bases de datos NoSQL
- 5) Técnicas y herramientas para extraer datos de la web
- 6) Implementar programas MapReduce en R con RHadoop
- 7) Técnicas para el procesado eficiente de conjuntos de datos a gran escala
- 8) Programación distribuida avanzada con Spark

### **Observaciones:**

Duración: 25 horas.

Fecha: 22, 24, 25 y 29 de noviembre y 1 de diciembre de 2016.

Horario: De 9:00 a 14:30 horas (descanso de 30 minutos incluido).

Prerrequisitos: Preferentemente los participantes deben poseer conocimientos de programación en R equivalentes a los adquiridos en el curso "Introducción a la plataforma estadística R".