

Actividad formativa específica (curso 2020-21)

Ciencia de datos en R

Ramas de conocimiento: orientado a doctorandos de la rama de Ciencias. En caso de que la demanda sea superior a las plazas ofertas, se priorizarán los estudiantes de los programa de doctorado del Departamento de Ciencias de la Vida.

Profesores: Sara Villén Pérez (coordinadora del curso; <https://saravillanperez.wixsite.com/villanperez>; sara.villen@uah.es) y Rafael Molina Venegas (<https://rafmolven.wixsite.com/rafmolven>; rafael.molina@uah.es). Área Ecología, Departamento de Ciencias de la Vida de la UAH.

Duración: 26 horas.

Fechas y horario de impartición: 31 de mayo – 4 junio.

Bloque I. 31 de mayo y 1 de junio; 11:00-14:00 & 16:00-18:30.

Bloque II. 2 y 3 de junio; 11:00-14:00 & 16:00-18:30.

Bloque II. 4 de junio; 10:00-14:00.

Lugar de impartición: Edificio de Ciencias de la UAH (presencial).

Número de plazas: 15 alumnos en cada bloque.

Inscripción: enviar solicitud por correo electrónico a pedro.villar@uah.es, especificando el programa de doctorado en el que está matriculado, el bloque o los bloques en los que desea inscribirse, así como una breve declaración del interés por el curso y de la experiencia previa en R. Se permitirá la inscripción en cada uno de los bloques del curso por separado.

Contenido:

Bloque I. Introducción y manejo de datos en R (11h; Prof. Sara Villén Pérez).

- Breve introducción a R.
- Estructuras y clases de datos.
- Importación y exportación, exploración, selección de subconjuntos, edición, reorganización, reestructuración y síntesis de bases de datos. Se verán funciones del paquete base y de la familia Tidyverse.

Bloque II. Visualización de datos con ggplot2 (11h; Prof. Sara Villén Pérez).

- Lógica de ggplot2 y gramática de los gráficos.
- Gráficos básicos: gráficos dispersión, de líneas, de área y de barras.
- Gráficos de distribución. Histogramas, polígonos de frecuencia y curvas de densidad; boxplots y gráficos de violín; mapas de frecuencia y de densidad.
- Visualización de datos espaciales: mapas.
- Gráficos animados.
- Apariencia general de los gráficos. Títulos, leyendas, ejes, fondo, etc.; ejes y gráficos circulares; visualización en paneles; paletas de colores; anotaciones. Impresión de figuras. Extensiones de ggplot2.

Bloque III. Automatización de procesos (4h; Prof. Rafael Molina Venegas).

- Bucles (if, if/else, while).
- Introducción a la programación de funciones.

Requerimientos: no es imprescindible (aunque sí recomendable) tener experiencia previa en R.