



# Curso Académico 2023-24

## SOFTWARE PARA GESTIÓN DE BASES DE DATOS

### Ficha Docente

#### ASIGNATURA

Nombre de asignatura (Código GeA): SOFTWARE PARA GESTIÓN DE BASES DE DATOS (608530)

Créditos: 6

Créditos presenciales: 1,80

Créditos no presenciales: 4,20

Semestre: 1

#### PLAN/ES DONDE SE IMPARTE

**Titulación:** BIOESTADÍSTICA  
**Plan:** MÁSTER UNIVERSITARIO EN BIOESTADÍSTICA  
**Curso:** 1 **Ciclo:** 2  
**Carácter:** OBLIGATORIA  
**Duración/es:** Primer cuatrimestre (actas en Feb. y Jul.)  
**Idioma/s en que se imparte:**  
**Módulo/Materia:** /

#### PROFESOR COORDINADOR

Nombre	Departamento	Centro	Correo electrónico	Teléfono
ALVAREZ LIEBANA, JAVIER	Estadística y Ciencia de los Datos	Facultad de Estudios Estadísticos	javlv09@ucm.es	

#### PROFESORADO

Nombre	Departamento	Centro	Correo electrónico	Teléfono
ALVAREZ LIEBANA, JAVIER	Estadística y Ciencia de los Datos	Facultad de Estudios Estadísticos	javlv09@ucm.es	

#### SINOPSIS

##### BREVE DESCRIPTOR:

Los estudiantes aprenderán a modelizar, construir y diseñar bases de datos multidimensionales de tal manera que se agilicen sus consultas y puedan ser explotadas posteriormente según el objetivo del estudio. También comprenderán la necesidad de estas técnicas, sus objetivos y sus aplicaciones en función del tipo de información disponible en cada caso.

##### REQUISITOS:

No hay requisitos previos.

##### OBJETIVOS:

##### COMPETENCIAS:

##### Generales:

CG3,CG5

##### Transversales:

##### Específicas:

CE6, CE7

##### Otras:

##### CONTENIDOS TEMÁTICOS:

1. Introducción a R y RStudio. Organización del trabajo. Introducción a R base. Tipos de variables y concatenación (vectores). Operaciones con vectores. Tipos de datos especiales: fechas y cadenas de textos.
2. Presentación de resultados: introducción a Quarto y rmarkdown. Elaborar apuntes y diapositivas.
3. Primeras estructuras de datos: matrices. Inconvenientes y limitaciones. Primeros datasets: data.frame y tibble.
4. Introducción a tidyverse. Tidydata. Pivotar tablas. Importar/exportar datasets.



# Curso Académico 2023-24

## SOFTWARE PARA GESTIÓN DE BASES DE DATOS

### Ficha Docente

5. Manejo de tidyverse. Operaciones por filas y columnas. Cruce de tablas. Depuración de datos: detección de ausentes y outliers e imputación. Recategorización de variables.
6. Estructuras de control, bucles y funciones propias. Variables locales vs globales. Manejo de listas y variables cualitativas. Paquetes purrr y forcats.
7. Introducción a módulos de acceso a SQL desde R.
8. Introducción a ggplot. Visualización de datos.
9. Control de versiones: uso de Github.
10. Introducción a Stata

#### ACTIVIDADES DOCENTES:

##### Clases teóricas:

30%

##### Clases prácticas:

70%

##### Trabajos de campo:

##### Prácticas clínicas:

##### Laboratorios:

##### Exposiciones:

##### Presentaciones:

##### Otras actividades:

##### TOTAL:

100%

#### EVALUACIÓN

A lo largo del curso se plantearán una serie de entregas individuales aplicando los conocimientos aprendidos a datasets reales. Se realizará también una entrega final grupal (máximo 4 personas) en la que tendrán que exponer el análisis realizado con un caso real.

La nota final en cualquier caso tendrá en cuenta tanto la evaluación continua como la prueba final y se calculará como el máximo entre:

a) Calificación de la prueba final.

b) La media ponderada de la evaluación continua y la prueba final, siendo el peso de la evaluación continua de al menos el 35% (pudiendo ser el 100%)

Cualquier alumno podrá presentarse al examen final, siendo la valoración del mismo el 100% de su nota final.

#### BIBLIOGRAFÍA BÁSICA:

- Wickham, H., Golemund, G.. «R for Data Science». Disponible gratuitamente en <https://r4ds.had.co.nz/index.html>
- Álvarez-Liébana, J. (2023). Curso introductorio de manejo y depuración de datos. «Manejo de datos en R». Disponible gratuitamente en <https://javieralvarezliebana.es/curso-R-RTVE/#/title-slide>
- Álvarez-Liébana, J. (2022). Curso introductorio a la visualización de datos. «Analizando datos, visualizando información, contando historias». Disponible gratuitamente en <https://javieralvarezliebana.es/curso-slides-dataviz-ECI-2022/#1>
- Álvarez-Liébana, J. (2021). Manual introductorio a R base. «Aprendiendo R sin morir en el intento». Disponible gratuitamente en <https://javieralvarezliebana.es/curso-intro-R/>
- Wright, C., Ellis, S.E., Hicks, S. C., Peng, R. D. (2021). «Tidyverse skills for data science». Disponible gratuitamente en <https://jhudatascience.org/tidyversecourse/>
- Matloff, N (2011). The Art of R Programming. A tour of statistical software design. Disponible como recurso electrónico en la biblioteca de la UCM.
- Crawley, M.J. The R book. Disponible como recurso electrónico en la biblioteca de la UCM

#### OTRA INFORMACIÓN RELEVANTE:

La presencialidad de las asignaturas de Máster es del 30%, lo que significa que la carga de trabajo de un alumno, para cada asignatura de 6 créditos ECTS, se divide en 45 horas lectivas presenciales y en 105 de trabajo personal adicional.



# Curso Académico 2023-24

## SOFTWARE PARA GESTIÓN DE BASES DE DATOS

### Ficha Docente