

## Guía docente de Asignatura – Grado en Ciencia de los Datos Aplicada

### Datos generales de la asignatura

<b>Asignatura:</b>	<b>Software Estadístico II - 806308</b>
<b>Curso académico:</b>	2023-2024
<b>Carácter</b>	Obligatoria
<b>Curso:</b>	Segundo
<b>Semestre:</b>	3
<b>Créditos ECTS</b>	
<b>Presenciales:</b>	2,4
<b>No presenciales:</b>	3,6
<b>Total</b>	6,0

### Actividades docentes

<b>Clases teóricas:</b>	
<b>Clases prácticas:</b>	
<b>Laboratorios:</b>	
<b>Total:</b>	
<b>Departamentos responsables:</b>	Estadística y Ciencia de los Datos
<b>Profesores:</b>	Javier Alvarez Liebana
<b>Profesor Coordinador:</b>	Javier Alvarez Liebana

### Datos específicos de la asignatura

<b>Breve descriptor:</b>	
<b>Requisitos:</b>	

### Competencias

<b>Generales:</b>	CG1, CG2, CG3, CG5, CG6, CG8, CG9
<b>Transversales:</b>	
<b>Específicas:</b>	CE7, CE10, CE11

### Objetivos

### Contenidos

1. Introducción a R y Rstudio. Organización del trabajo. Manejo de errores.
2. Introducción a R base. Tipos de variables y concatenación (vectores). Operaciones con vectores. Tipos de datos especiales: fechas y cadenas de textos. Uso de primeros paquetes.
3. Presentación de resultados: introducción a Quarto y Rmarkdown. Elaborar apuntes y diapositivas.
4. Primera estructuras de datos: matrices. Inconvenientes y limitaciones. Primeros datasets: data.frame y tibble. Ventajas respecto a data.frame: rapidez, legibilidad, recursividad y memoria.
5. Introducción a tidyverse. Tidydata. Pivotar tablas. Importar/exportar datasets. Cruce de tablas.
6. Manejo de tidyverse. Operaciones por filas y columnas.
7. Estructuras de control, bucles y funciones propias. Variables locales vs globales.
8. Manejo de listas y cualitativas. Paquetes purrr y forcats.
9. Introducción a ggplot. Visualización de datos. Animación de gráficas y mapas.
10. Uso de control de versiones. Github y gitkraken.

### 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 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

- Wickham, H., Golemund, G.. «R for Data Science». Disponible gratuitamente en <https://r4ds.had.co.nz/index.html>
- Álvarez-Liébaña, 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>
- Alvarez-Liébaña, 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>
- Alvarez-Liébaña, 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