

Facultad de Ciencias

GUÍA DOCENTE DE LA ASIGNATURA

M1965 - Estadística para la Ciencia de Datos

Máster Universitario en Ciencia de Datos / Master in Data Science
Obligatoria. Curso 1

Curso Académico 2021-2022

1. DATOS IDENTIFICATIVOS

Título/s	Máster Universitario en Ciencia de Datos / Master in Data Science	Tipología y Curso	Obligatoria. Curso 1
Centro	Facultad de Ciencias		
Módulo / materia	FUNDAMENTOS MÉTODOS EN CIENCIA DE DATOS		
Código y denominación	M1965 - Estadística para la Ciencia de Datos		
Créditos ECTS	6	Cuatrimestre	Cuatrimestral (1)
Web			
Idioma de impartición	Español	English friendly	Sí
		Forma de impartición	Presencial

Departamento	DPTO. MATEMATICAS, ESTADISTICA Y COMPUTACION
Profesor responsable	MARCOS CRUZ RODRIGUEZ
E-mail	marcos.cruz@unican.es
Número despacho	Facultad de Ciencias. Planta: + 1. DESPACHO (1015)
Otros profesores	PABLO MARTINEZ RUIZ DEL ARBOL ANA CASANUEVA VICENTE DANIEL GARCIA DIAZ

2. CONOCIMIENTOS PREVIOS

Conocimientos de estadística básica y programación.

3. COMPETENCIAS GENÉRICAS Y ESPECÍFICAS DEL PLAN DE ESTUDIOS TRABAJADAS

Competencias Genéricas
Capacidad de estudio, síntesis y autonomía suficientes para desarrollar de forma autónoma proyectos básicos de investigación.
Buscar, obtener, procesar, comunicar información y transformarla en conocimiento.
Conocer las herramientas metodológicas necesarias para desarrollar proyectos avanzados.
Integrarse eficazmente en un grupo de trabajo y trabajar en equipo, compartir la información disponible e integrar su actividad en la actividad del grupo colaborando de forma activa en la consecución de objetivos comunes.
Capacidad de actualización de los conocimientos expuestos en el ámbito de la comunidad científica.
Competencias Específicas
Utilizar el análisis predictivo para analizar grandes volúmenes de datos y descubrir nuevas relaciones.
Utilizar técnicas estadísticas apropiadas sobre los datos disponibles para lograr una visión adecuada de los mismos.
Aplicar el ingenio propio para resolver problemas complejos y desarrollar ideas innovadoras.
Competencias Básicas
Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio.
Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios.
Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades.
Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.
Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación.
Competencias Transversales
Analizar y combinar información utilizando diferentes fuentes.
Dominio de la gestión del tiempo.
Capacidad de trabajo autónomo y toma de decisiones.
Capacidades asociadas al trabajo en equipo: cooperación, liderazgo, saber escuchar.

3.1 RESULTADOS DE APRENDIZAJE

- Saber aplicar métodos de estadística descriptiva
- Saber aplicar métodos de muestreo y Monte Carlo.
- Saber los fundamentos de inferencia estadística
- Saber aplicar los contrastes paramétricos y no paramétricos para una y varias muestras.
- Saber aplicar técnicas de remuestreo (bootstrap).
- Saber aplicar modelos de regresión y estimación de máxima verosimilitud.
- Saber aplicar regularización y regresión contraída (ridge regression)

4. OBJETIVOS

Conocer los métodos de estadística descriptiva
Conocer técnicas de muestreo y Monte Carlo.
Conocer los fundamentos de inferencia estadística.
Conocer los contrastes paramétricos y no paramétricos para una y varias muestras.
Conocer las técnicas de remuestreo (bootstrap).
Conocer los modelos de regresión y estimación de máxima verosimilitud.
Conocer los métodos de regularización y regresión contraída (ridge regression)

5. MODALIDADES ORGANIZATIVAS Y MÉTODOS DOCENTES

ACTIVIDADES	HORAS DE LA ASIGNATURA
ACTIVIDADES PRESENCIALES	
HORAS DE CLASE (A)	
- Teoría (TE)	30
- Prácticas en Aula (PA)	
- Prácticas de Laboratorio Experimental(PLE)	
- Prácticas de Laboratorio en Ordenador (PLO)	30
- Prácticas Clínicas (CL)	
Subtotal horas de clase	60
ACTIVIDADES DE SEGUIMIENTO (B)	
- Tutorías (TU)	7
- Evaluación (EV)	3
Subtotal actividades de seguimiento	10
Total actividades presenciales (A+B)	70
ACTIVIDADES NO PRESENCIALES	
Trabajo en grupo (TG)	
Trabajo autónomo (TA)	80
Tutorías No Presenciales (TU-NP)	
Evaluación No Presencial (EV-NP)	
Total actividades no presenciales	80
HORAS TOTALES	150

6. ORGANIZACIÓN DOCENTE													
CONTENIDOS		TE	PA	PLE	PLO	CL	TU	EV	TG	TA	TU-NP	EV-NP	Semana
1	1. Estadística descriptiva. 2. Muestreo y Monte Carlo. 3. Fundamentos de inferencia estadística: una y varias variables. 4. Contrastes paramétricos y no paramétricos para una y varias muestras.	12,50	0,00	0,00	12,50	0,00	3,00	1,00	0,00	30,00	0,00	0,00	1-3
2	5. Técnicas de remuestreo (bootstrap). 6. Modelos de regresión. Estimación de máxima verosimilitud. 7. Regularización. Regresión contraída (ridge regression)	17,50	0,00	0,00	17,50	0,00	4,00	2,00	0,00	50,00	0,00	0,00	4-7
TOTAL DE HORAS		30,00	0,00	0,00	30,00	0,00	7,00	3,00	0,00	80,00	0,00	0,00	
Esta organización tiene carácter orientativo.													

Ante la situación incierta de que las medidas de distanciamiento social establecidas por las autoridades sanitarias no permitan desarrollar alguna actividad docente de forma presencial en el aula para todos los estudiantes matriculados, se adoptará una modalidad mixta de docencia que combine esta docencia presencial en el aula con docencia a distancia. De la misma manera, la tutorización podrá ser sustituida por tutorización a distancia utilizando medios telemáticos.

TE	Horas de teoría
PA	Horas de prácticas en aula
PLE	Horas de prácticas de laboratorio experimental
PLO	Horas de prácticas de laboratorio en ordenador
CL	Horas de prácticas clínicas
TU	Horas de tutoría
EV	Horas de evaluación
TG	Horas de trabajo en grupo
TA	Horas de trabajo autónomo
TU-NP	Tutorías No Presenciales
EV-NP	Evaluación No Presencial

7. MÉTODOS DE LA EVALUACIÓN

Descripción	Tipología	Eval. Final	Recuper.	%										
Test	Examen escrito	Sí	Sí	40,00										
<table border="1"> <tr> <td>Calif. mínima</td> <td>0,00</td> </tr> <tr> <td>Duración</td> <td></td> </tr> <tr> <td>Fecha realización</td> <td>Al final del curso</td> </tr> <tr> <td>Condiciones recuperación</td> <td></td> </tr> <tr> <td>Observaciones</td> <td></td> </tr> </table>		Calif. mínima	0,00	Duración		Fecha realización	Al final del curso	Condiciones recuperación		Observaciones				
Calif. mínima	0,00													
Duración														
Fecha realización	Al final del curso													
Condiciones recuperación														
Observaciones														
Valoración de informes y trabajos escritos	Trabajo	Sí	Sí	60,00										
<table border="1"> <tr> <td>Calif. mínima</td> <td>3,00</td> </tr> <tr> <td>Duración</td> <td></td> </tr> <tr> <td>Fecha realización</td> <td>Durante el periodo de impartición de la asignatura</td> </tr> <tr> <td>Condiciones recuperación</td> <td></td> </tr> <tr> <td>Observaciones</td> <td></td> </tr> </table>		Calif. mínima	3,00	Duración		Fecha realización	Durante el periodo de impartición de la asignatura	Condiciones recuperación		Observaciones				
Calif. mínima	3,00													
Duración														
Fecha realización	Durante el periodo de impartición de la asignatura													
Condiciones recuperación														
Observaciones														
TOTAL				100,00										
Observaciones														
Criterios de evaluación para estudiantes a tiempo parcial														
Los alumnos a tiempo parcial realizarán las mismas pruebas que los demás alumnos.														

8. BIBLIOGRAFÍA Y MATERIALES DIDÁCTICOS

BÁSICA
María Dolores Ugarte, Ana F. Militino, and Alan T. Arnholt: "Probability and Statistics with R", Second Edition. Chapman & Hall 2015.
Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani: "An Introduction to Statistical Learning" (with Applications in R). Springer-Verlag 2013.
Complementaria
Hastie, Tibshirani and Friedman: "The Elements of Statistical Learning" (2nd edition). Springer-Verlag 2009.
David MacKay: "Information Theory, Inference, and Learning Algorithms" Cambridge University Press, 2003.

9. SOFTWARE

PROGRAMA / APLICACIÓN	CENTRO	PLANTA	SALA	HORARIO
Software R				
Python				

10. COMPETENCIAS LINGÜÍSTICAS

- | | |
|---|---|
| <input checked="" type="checkbox"/> Comprensión escrita | <input type="checkbox"/> Comprensión oral |
| <input type="checkbox"/> Expresión escrita | <input type="checkbox"/> Expresión oral |
| <input type="checkbox"/> Asignatura íntegramente desarrollada en inglés | |

Observaciones