



Guía docente				
Datos Identificativos				2019/20
Asignatura (*)	Tecnologías de Gestión de Datos		Código	614493106
Titulación	Mestrado Universitario en Técnicas Estadísticas (Plan 2019)			
Descriptores				
Ciclo	Periodo	Curso	Tipo	Créditos
Máster Oficial	1º cuatrimestre	Primero	Optativa	5
Idioma	Castellano			
Modalidad docente	Presencial			
Prerrequisitos				
Departamento	Enxeñaría de ComputadoresMatemáticas			
Coordinador/a	López Taboada, Guillermo	Correo electrónico	guillermo.lopez.taboada@udc.es	
Profesorado	Fernández Casal, Rubén	Correo electrónico	ruben.fcasal@udc.es	
	López Taboada, Guillermo		guillermo.lopez.taboada@udc.es	
Web	eio.usc.es/pub/mte			
Descripción general	El objetivo de esta asignatura es que cualquier alumno, independientemente de su formación previa, adquiera un sólido conocimiento de las tecnologías de gestión de bases de datos, tanto relacionales como no relacionales. Asimismo, se buscará la familiarización con las principales técnicas computacionales para la gestión práctica de datos masivos. Esto dotará al alumno de una gran autonomía a la hora de procesar y estudiar datos, independientemente de su formato y origen.			

Competencias del título	
Código	Competencias del título
A16	CE1 - Conocer, identificar, modelar, estudiar y resolver problemas complejos de estadística e investigación operativa, en un contexto científico, tecnológico o profesional, surgidos en aplicaciones reales
A17	CE2 - Desarrollar autonomía para la resolución práctica de problemas complejos surgidos en aplicaciones reales y para la interpretación de los resultados de cara a la ayuda en la toma de decisiones.
A18	CE3 - Adquirir conocimientos avanzados de los fundamentos teóricos subyacentes a las distintas metodologías de la estadística y la investigación operativa, que permitan su desarrollo profesional especializado.
A21	CE6 - Adquirir conocimientos teórico-prácticos avanzados de distintas técnicas matemáticas, orientadas específicamente a la ayuda en la toma de decisiones, y desarrollar la capacidad de reflexión para evaluar y decidir entre distintas perspectivas en contextos complejos.
A23	CE8 - Adquirir conocimientos teórico-prácticos avanzados de las técnicas destinadas a la realización de inferencias y contrastes relativos a variables y parámetros de un modelo estadístico, y saber aplicarlos con autonomía suficiente un contexto científico, tecnológico o profesional.
A24	CE9 - Conocer y saber aplicar con autonomía en contextos científicos, tecnológicos o profesionales, técnicas de aprendizaje automático y técnicas de análisis de datos de alta dimensión (big data).
A25	CE10 - Adquirir conocimientos avanzados sobre metodologías para la obtención y el tratamiento de datos desde distintas fuentes, como encuestas, internet, o entornos "en la nube";.
B1	CB6 - 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
B2	CB7 - 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
B3	CB8 - 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
B4	CB9 - 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
B5	CB10 - 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.



B17	CG1 - Conocer, comprender y saber aplicar los principios, metodologías y nuevas tecnologías en la estadística y la investigación operativa en contextos científico/académicos, tecnológicos o profesionales especializados y multidisciplinares, así como adquirir las destrezas y competencias descritas en los objetivos generales del título.
B18	CG2 - Desarrollar autonomía para identificar, modelar y resolver problemas complejos de la estadística y la investigación operativa en contextos científico/académicos, tecnológicos o profesionales especializados y multidisciplinares.
B19	CG3 - Desarrollar la capacidad para realizar estudios y tareas de investigación y transmitir los resultados a públicos especializados, académicos y generalistas.
B20	CG4 - Integrar conocimientos avanzados y enfrentarse a la toma de decisiones a partir de información científica y técnica.
B21	CG5 - Desarrollar la capacidad de aplicación de algoritmos y técnicas de resolución de problemas complejos en el ámbito de la estadística y la investigación operativa, manejando el software especializado adecuado.
C11	CT1 - Desarrollar firmes capacidades de razonamiento, análisis crítico y autocrítico, así como de argumentación y de síntesis, contextos especializados y multidisciplinares.
C12	CT2 - Desarrollar destrezas avanzadas en el manejo de Tecnologías de la Información y la Comunicación (TIC), tanto para la obtención de información como para la difusión del conocimiento, en un ámbito científico/académico, tecnológico o profesional especializado y multidisciplinar.
C13	CT3 - Ser capaz de resolver problemas complejos en entornos nuevos mediante la aplicación integrada de los conocimientos.
C14	CT4 - Desarrollar una sólida capacidad de organización y planificación del estudio, asumiendo la responsabilidad de su propio desarrollo profesional, para la realización de trabajos en equipo y de forma autónoma.
C15	CT5 - Desarrollar capacidades para el aprendizaje y la integración en el trabajo en equipos multidisciplinares, en los ámbitos científico/académico, tecnológico y profesional.

Resultados de aprendizaje			
Resultados de aprendizaje	Competencias del título		
Manejar de forma autónoma y solvente el software necesario para acceder a conjuntos de datos en entornos profesionales y/o en la nube.	AM16 AM17 AM23 AM24 AM25	BP5 BP17 BP19 BP20 BP21	CP12
Saber gestionar conjuntos de datos masivos en un entorno multidisciplinar que permita la participación en proyectos profesionales complejos que requieran el uso de técnicas estadísticas.	AM18 AM21	BP1 BP2 BP3 BP4 BP18	CP11 CP13 CP14 CP15
Saber relacionar el software de diseño y gestión de bases de datos con el específicamente implementado para el análisis de datos.	AM16 AM17 AM21 AM24 AM25	BP17 BP18 BP21	CP12 CP13

Contenidos	
Tema	Subtema
1. Introducción al lenguaje SQL	1.1 Bases de datos relacionales 1.2 Sintaxis SQL 1.3 Conexión con bases de datos desde R
2. Introducción a tecnologías NoSQL	2.1 Conceptos y tipos de bases de datos NoSQL (documental, columnar, clave/valor y de grafos) 2.2 Conexión de R a NoSQL



3. Tecnologías para el tratamiento de datos masivos	3.1 Tecnologías Big Data (Hadoop, Spark, Hive, Rspark, Sparklyr) 3.2 Visualización y generación de cuadros de mando 3.3 Introducción al análisis de datos masivos.
---	--

Planificación				
Metodologías / pruebas	Competencias	Horas presenciales	Horas no presenciales / trabajo autónomo	Horas totales
Sesión magistral	A16 A18 A21 A23	20	28	48
Prácticas de laboratorio	A16 A18 A24 A25 B3 B2 C13 C12	7	10.5	17.5
Solución de problemas	A17 A16 B2 C13	0	28	28
Seminario	A17 A24 A25 B1 B17 B19 B21 C15	7	10.5	17.5
Trabajos tutelados	A16 A17 A21 A23 A24 A25 B1 B2 B3 B4 B5 B17 B18 B19 B20 B21 C11 C12 C13 C14 C15	1	9.5	10.5
Prueba mixta	A16 A18 A21 A23 A24 A25 B1 B2 B3 B4 B17 B18 C13	3	0	3
Atención personalizada		0.5	0	0.5
(*)Los datos que aparecen en la tabla de planificación són de carácter orientativo, considerando la heterogeneidad de los alumnos				

Metodologías	
Metodologías	Descripción
Sesión magistral	Sesiones expositivas, en las que los presentarán conceptos y/o procedimientos, aportando información básica necesaria para entender una perspectiva teórica o un procedimiento práctico, promoviendo la participación del estudiantado.
Prácticas de laboratorio	Sesiones interactivas de prácticas de laboratorio (informática) o de resolución de problemas, donde los docentes apoyarán y supervisarán la puesta en práctica de los conocimientos adquiridos por parte del alumnado.
Solución de problemas	Trabajo personal del alumnado en la realización de los ejercicios prácticos y resolución de problemas a partir de las sesiones magistrales y prácticas de laboratorio.
Seminario	Técnica de trabajo en grupo que tiene como finalidad el estudio intensivo de un tema. Se caracteriza por la discusión, la participación, la elaboración de documentos y las conclusiones a las que tienen que llegar todos los componentes del seminario.
Trabajos tutelados	Actividades de aprendizaje autónomo, en las que los docentes guiarán la realización de trabajos individuales o en pequeños grupos por parte del alumnado.
Prueba mixta	Prueba que integra preguntas tipo de pruebas de ensayo y preguntas tipo de pruebas objetivas. En cuanto a preguntas de ensayo, recoge preguntas abiertas de desarrollo. Además, en cuanto preguntas objetivas, puede combinar preguntas de respuesta múltiple, de ordenación, de respuesta breve, de discriminación, de completar y/o de asociación.

Atención personalizada	
Metodologías	Descripción



Prácticas de laboratorio Seminario Solución de problemas Trabajos tutelados	Atención personalizada (o en grupo reducidos/muy reducidos) en estas metodologías, tanto en el aula como en horario de tutorías, para solventar las dudas que puedan surgir en el desarrollo de la docencia y que no hayan sido resueltas de forma eficaz previamente.
--	--

Evaluación			
Metodologías	Competencias	Descripción	Calificación
Prácticas de laboratorio	A16 A18 A24 A25 B3 B2 C13 C12	Evaluación de las prácticas de laboratorio desarrolladas por los estudiantes.	30
Prueba mixta	A16 A18 A21 A23 A24 A25 B1 B2 B3 B4 B17 B18 C13	El examen de la materia evaluará los siguientes aspectos: Conceptos de la materia: Dominio de los conocimientos teóricos y operativos de la materia. Asimilación práctica de materia: Asimilación y comprensión de los conocimientos teóricos y operativos de la materia.	60
Trabajos tutelados	A16 A17 A21 A23 A24 A25 B1 B2 B3 B4 B5 B17 B18 B19 B20 B21 C11 C12 C13 C14 C15	Evaluación de los trabajos tutelados desarrollados por los estudiantes.	10

Observaciones evaluación
Para poder aprobar la asignatura en la primera oportunidad será necesario obtener como mínimo el 30% de la nota máxima de la suma de las prácticas de laboratorio y trabajos tutelados e, igualmente, el 30% de la nota máxima final de la Prueba mixta (examen), y tener una nota total (prácticas más trabajos tutelados más prueba mixta) igual o superior al 50% de la nota máxima. En la segunda oportunidad solamente se podrá recuperar la nota del examen. Las notas de prácticas y de trabajos tutelados serán las obtenidas durante el curso. Para los alumnos que utilicen la oportunidad adelantada de diciembre se utilizarán las notas de prácticas y trabajos tutelados que obtuvieran en su último curso. En esta oportunidad solo será necesario para aprobar obtener una nota total igual o superior al 50% de la nota máxima. Una vez que un estudiante es evaluado en una práctica de laboratorio o en un trabajo tutelado implica que será calificado. Por tanto, la calificación "No Presentado" no es posible una vez que una práctica/trabajo ha sido evaluada.

Fuentes de información	
Básica	- Daroczi, G. (2015). Mastering Data Analysis with R. Packt Publishing - Grolemond, G. y Wickham, H. (2016). R for Data Science. https://r4ds.had.co.nz/ & O'Reilly - Silberschatz, A., Korth, H. y Sudarshan, S. (2014). Fundamentos de Bases de Datos. Mc Graw Hill - Rubén Fernández Casal (2019). Ayuda y Recursos para el Aprendizaje de R. https://rubenfcasal.github.io/post/ayuda-y-recursos-para-el-aprendizaje-de-r/
Complementaria	- Wes McKinney (2017). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. O'Reilly (2ª ed.) - Tom White (2015). Hadoop: The Definitive Guide. O'Reilly (4ª ed.) - Alex Holmes (2014). Hadoop in practice. Manning (2ª ed.) - Centro de Supercomputación de Galicia (2019). Servicio de Big Data del CESGA. https://bigdata.cesga.es/

Recomendaciones
Asignaturas que se recomienda haber cursado previamente
Asignaturas que se recomienda cursar simultáneamente
Asignaturas que continúan el temario



Otros comentarios

Recomendaciones para el estudio de la materia Debido al fuerte componente práctico es recomendable ir haciendo las actividades prácticas y trabajos académicamente dirigidos de forma regular a lo largo del cuatrimestre. Las herramientas software utilizadas en esta materia son generalmente open-source o tienen licencia gratuita para estudiantes.

(*) La Guía Docente es el documento donde se visualiza la propuesta académica de la UDC. Este documento es público y no se puede modificar, salvo cosas excepcionales bajo la revisión del órgano competente de acuerdo a la normativa vigente que establece el proceso de elaboración de guías