



Guía docente				
Datos Identificativos				2022/23
Asignatura (*)	Tratamiento de Datos Masivos	Código	710G04035	
Titulación	Grao en Xestión Dixital de Información e Documentación			
Descritores				
Ciclo	Periodo	Curso	Tipo	Créditos
Grado	1º cuatrimestre	Cuarto	Optativa	6
Idioma	Castellano			
Modalidad docente	Presencial			
Prerrequisitos				
Departamento	Ciencias da Computación e Tecnoloxías da Información			
Coordinador/a	Cortiñas Álvarez, Alejandro	Correo electrónico	alejandro.cortinas@udc.es	
Profesorado	Cortiñas Álvarez, Alejandro González Rueda, Ángel Manuel	Correo electrónico	alejandro.cortinas@udc.es angel.manuel.rueda@udc.es	
Web				
Descripción general	El objetivo de la asignatura es desarrollar conocimientos teóricos y prácticos sobre las tecnologías NoSQL y BigData, así como de análisis y modelado estadístico.			

Competencias / Resultados del título	
Código	Competencias / Resultados del título
A1	CE1 - Conocer y comprender los principios teóricos y metodológicos de la gestión de información y la documentación para aplicarlos en su actividad profesional
A6	CE6 - Buscar y recuperar información en diversos medios para dar respuesta a la demanda de los usuarios de información
A7	CE7 - Planificar y diseñar un sistema de gestión de la información, incluyendo los flujos de información, tanto en un contexto institucional como empresarial
A8	CE8 - Dominar los diferentes métodos de representación de los datos, información y el conocimiento que garanticen su recuperación eficiente
A13	CE13 - Conocer y dominar las técnicas y normativas para la creación y autenticación, reunión, selección, organización, representación, preservación, recuperación, acceso, difusión e intercambio, y evaluación de los recursos y servicios de información
A20	CE20 - Dominar las bases para desarrollar actividades de investigación utilizando métodos y principios multidisciplinares
A21	CE21 - Poseer conocimientos de estadística y análisis cuantitativo de la información
A22	CE22 - Adquirir habilidades computacionales y de manejo de las nuevas TIC

Resultados de aprendizaje			
Resultados de aprendizaje		Competencias / Resultados del título	
Identificar y analizar las situaciones en las que se puede aplicar la tecnología Big Data.		A1	
		A6	
		A7	
		A8	
		A13	
Adquirir habilidades para realizar conexiones con bases de datos relacionales desde software de referencia en análisis de datos.		A1	
		A7	
		A8	
		A13	
		A22	



Entender los conceptos y conocer los tipos de bases de datos NoSQL (documental, columnar, clave/valor, de grafos)	A1 A7 A8 A13 A22		
Adquirir habilidades para realizar conexiones a NoSQL desde software de referencia en análisis de datos.	A1 A7 A8 A22		
Conocer las principales tecnologías Big Data como son Hadoop, Spark, Hive, Rspark, y Sparklyr, entre otras.	A1 A6 A7 A8 A13		
Capacidades para la visualización y generación de cuadros de mando (por ejemplo con shiny)	A1 A7 A8 A22		
Ser capaz de aplicar técnicas de análisis estadístico de datos masivos.	A1 A7 A8 A21		
Desenvolver conocimientos de análisis de datos computacional, incluyendo programas como R.	A1 A7 A8 A21		
Capacidad de análisis y de síntesis aplicada a la gestión y organización de la información.	A1 A7 A8 A20 A21		
Adquisición de habilidades para la toma de decisiones a partir del análisis estadístico de bases de datos complejas.	A20 A21		

Contenidos	
Tema	Subtema
NoSQL y BigData	Introducción a tecnologías NoSQL. Tecnologías Big Data. Visualización y generación de cuadros de mando.
Análisis y modelos estadísticos	Introducción al análisis estadístico de datos masivos. Generación de informes estadísticos dinámicos y cuadros de mando con R para la gestión de datos. Modelos de regresión para datos de alta dimensión. Modelos de clasificación para datos de alta dimensión.

Planificación				
Metodologías / pruebas	Competencias / Resultados	Horas lectivas (presenciales y virtuales)	Horas trabajo autónomo	Horas totales



Prácticas a través de TIC	A1 A6 A7 A8 A13 A20 A21 A22	13	0	13
Estudio de casos	A1 A6 A7 A8 A13 A20 A21 A22	7	0	7
Trabajos tutelados	A1 A6 A7 A8 A13 A20 A21 A22	0	101	101
Prueba objetiva	A1 A6 A7 A8 A13 A20 A21 A22	0	7	7
Sesión magistral	A1 A6 A7 A8 A13 A20 A21 A22	21	0	21
Atención personalizada		1	0	1

(*Los datos que aparecen en la tabla de planificación són de carácter orientativo, considerando la heterogeneidad de los alumnos

Metodologías	
Metodologías	Descripción
Prácticas a través de TIC	Metodología que permite al alumnado aprender de forma efectiva, a través de actividades de carácter práctico (demostraciones, simulaciones, etc.) la teoría de un ámbito de conocimiento, mediante la utilización de las tecnologías de la información y las comunicaciones. Las TIC suponen un excelente soporte y canal para el tratamiento de la información y aplicación práctica de conocimientos, facilitando el aprendizaje y el desarrollo de habilidades por parte del alumnado.
Estudio de casos	Metodología donde el sujeto se enfrenta ante la descripción de una situación específica que plantea un problema que ha de ser comprendido, valorado y resuelto por un grupo de personas, a través de un proceso de discusión. El alumno se sitúa ante un problema concreto (caso), que le describe una situación real de la vida profesional, y debe ser capaz de analizar una serie de hechos, referentes a un campo particular del conocimiento o de la acción, para llegar a una decisión razonada a través de un proceso de discusión en pequeños grupos de trabajo.
Trabajos tutelados	Metodología diseñada para promover el aprendizaje autónomo de los estudiantes, bajo la tutela del profesor y en escenarios variados (académicos y profesionales). Está referida prioritariamente al aprendizaje del ¿cómo hacer las cosas?. Constituye una opción basada en la asunción por los estudiantes de la responsabilidad por su propio aprendizaje. Este sistema de enseñanza se basa en dos elementos básicos: el aprendizaje independiente de los estudiantes y el seguimiento de ese aprendizaje por el profesor-tutor.
Prueba objetiva	Prueba escrita utilizada para la evaluación del aprendizaje, cuyo rasgo distintivo es la posibilidad de determinar si las respuestas dadas son o no correctas. Constituye un instrumento de medida, elaborado rigurosamente, que permite evaluar conocimientos, capacidades, destrezas, rendimiento, aptitudes, actitudes, inteligencia, etc. Es de aplicación tanto para la evaluación diagnóstica, formativa como sumativa. La prueba objetiva puede combinar distintos tipos de preguntas: preguntas de respuesta múltiple, de ordenación, de respuesta breve, de discriminación, de completar y/o de asociación. También se puede construir con un solo tipo de alguna de estas preguntas.
Sesión magistral	Exposición oral complementada con el uso de medios audiovisuales y la introducción de algunas preguntas dirigidas a los estudiantes, con la finalidad de transmitir conocimientos y facilitar el aprendizaje. La clase magistral es también conocida como ¿conferencia?, ¿método expositivo? o ¿lección magistral?. Esta última modalidad se suele reservar a un tipo especial de lección impartida por un profesor en ocasiones especiales, con un contenido que supone una elaboración original y basada en el uso casi exclusivo de la palabra como vía de transmisión de la información a la audiencia.

Atención personalizada	
Metodologías	Descripción



Trabajos tutelados	Se estima que entre el alumnado habrá diferencias notables tanto en cuanto a su familiarización con conceptos y términos informáticos, como en cuanto a las habilidades para el manejo de herramientas informáticas. Por eso, se prevé desarrollar una atención personalizada para los trabajos tutelados. La atención personalizada se desarrollará de forma individual durante las clases o en horas de tutoría.
--------------------	--

Evaluación			
Metodologías	Competencias / Resultados	Descripción	Calificación
Prueba objetiva	A1 A6 A7 A8 A13 A20 A21 A22	Prueba individual de contenidos teórico-prácticos	40
Trabajos tutelados	A1 A6 A7 A8 A13 A20 A21 A22	Se valorará la calidad de dos trabajos realizados	60

Observaciones evaluación
<p>PRIMERA OPORTUNIDAD</p> <p>Para aprobar la materia es obligatorio:</p> <p>Una NOTA MÍNIMA de 3 (sobre 6) en los trabajos tutelados. Una NOTA MÍNIMA de 2 (sobre 4) en la prueba objetiva. De no obtener la nota mínima en los trabajos tutelados o en la prueba objetiva, la nota máxima global de la materia no será superior a un 4,5.</p> <p>Tendrá calificación de NO PRESENTADO cualquier estudiante que no realice la prueba objetiva.</p> <p>SEGUNDA OPORTUNIDAD</p> <p>Podrán presentarse a la segunda oportunidad ÚNICAMENTE aquellos/as estudiantes que no superen la materia en la primera oportunidad. Las condiciones son las mismas que en la primera oportunidad, incluyendo la nota mínima en ambas partes. Si un/una estudiante decide no realizar la recuperación de alguna de las partes, conservará la nota obtenida en la primera oportunidad en esa parte.</p> <p>Tendrá calificación de NO PRESENTADO cualquier estudiante que no opte a la recuperación de ninguna dos partes.</p> <p>DISPENSA ACADÉMICA</p> <p>Dado que la asistencia a las sesiones presenciales no es obligatoria, aquellos/as estudiantes con matrícula a tiempo parcial y dispensa académica que les exima de la asistencia a las clases tendrán las mismas condiciones que el resto del alumnado.</p> <p>OPORTUNIDAD ADELANTADA</p> <p>Se utilizarán los mismos criterios que en las oportunidad normales.</p> <p>IMPLICACIONES DEL PLAGIO</p> <p>La realización fraudulenta de pruebas o actividades de evaluación, una vez verificadas, implicará directamente una calificación de "0" en la asignatura en la oportunidad adecuada.</p>

Fuentes de información	
Básica	<ul style="list-style-type: none"> - James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: springer - Campbell, M. (2019). Learn RStudio IDE (pp. 39-48). <p>Dado o caracter esencialmente práctico desta asignatura, a bibliografía estará principalmente composta polos manuais das ferramentas informáticas descritas.</p>
Complementaria	

Recomendaciones
Asignaturas que se recomienda haber cursado previamente
Lenguajes de Programación, Análisis y Consulta para la Gestión de Información/710G04029
Sistemas de Información para la Gestión Documental/710G04025
Informática para el Tratamiento y Gestión de la Información/710G04024
Asignaturas que se recomienda cursar simultáneamente
Asignaturas que continúan el temario



Otros comentarios

(*) 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