



Guía Docente				
Datos Identificativos				2022/23
Asignatura (*)	Tratamento de Datos Masivos	Código	710G04035	
Titulación	Grao en Xestión Dixital de Información e Documentación			
Descritores				
Ciclo	Período	Curso	Tipo	Créditos
Grao	1º cuatrimestre	Cuarto	Optativa	6
Idioma	Castelán			
Modalidade docente	Presencial			
Prerrequisitos				
Departamento	Ciencias da Computación e Tecnoloxías da Información			
Coordinación	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				
Descrición xeral	O obxectivo do curso é desenvolver coñecementos teóricos e prácticos sobre as tecnoloxías NoSQL e BigData, así como a análise e modelado estatístico.			

Competencias / Resultados do título	
Código	Competencias / Resultados do título
A1	CE1 - Coñecer e comprender os principios teóricos e metodolóxicos da xestión e documentación da información para aplicalos na súa actividade profesional
A6	CE6 - Buscar e obter información en diversos medios para responder á demanda dos usuarios de información
A7	CE7 - Planificar e deseñar un sistema de xestión da información, incluíndo os fluxos de información, tanto nun contexto institucional como empresarial
A8	CE8 - Dominar os distintos métodos de representación de datos, información e coñecemento que garanta a súa recuperación eficiente
A13	CE13 - Coñecer e dominar as técnicas e normativas para a creación e autenticación, reunión, selección, organización, representación, preservación, recuperación, acceso, difusión, intercambio e avaliación de recursos e servizos de información
A20	CE20 - Dominar as bases para desenvolver actividades de investigación utilizando métodos e principios multidisciplinares
A21	CE21 - Posuír coñecementos de estatística e análise cuantitativa da información
A22	CE22 - Adquirir habilidades computacionais e xestión das novas TIC

Resultados da aprendizaxe			
Resultados de aprendizaxe	Competencias / Resultados do título		
	Identifica e analiza as situacións nas que se pode aplicar a tecnoloxía Big Data.	A1	
	A6		
	A7		
	A8		
	A13		
Adquirir habilidades para establecer conexións con bases de datos relacionais a partir de software de referencia na análise de datos.	A1		
	A7		
	A8		
	A13		
	A22		
Comprender os conceptos e coñecer os tipos de bases de datos NoSQL (documentais, columnares, clave/valor, gráfica)	A1		
	A7		
	A8		
	A13		
	A22		



Adquirir habilidades para establecer conexións con NoSQL a partir de software de referencia na análise de datos.	A1 A7 A8 A22		
Coñece as principais tecnoloxías de Big Data como Hadoop, Spark, Hive, Rspark e Sparklyr, entre outras.	A1 A6 A7 A8 A13		
Capacidade para a visualización e xeración de paneis (por exemplo con brillante)	A1 A7 A8 A22		
Ser capaz de aplicar técnicas de análise estatística a datos masivos.	A1 A7 A8 A21		
Desenvolver coñecementos de análise de datos computacionais, incluíndo programas como R.	A1 A7 A8 A21		
Capacidade de análise e síntese aplicada á xestión e organización da información.	A1 A7 A8 A20 A21		
Adquisición de habilidades para a toma de decisións a partir da análise estatística de bases de datos complexas.	A20 A21		

Contidos	
Temas	Subtemas
NoSQL e Big Data	Introdución ás tecnoloxías NoSQL. Tecnoloxías de Big Data. Visualización e xeración de paneis.
Análise estatística e modelos	Introdución á análise estatística de datos masivos. Xeración de informes estatísticos dinámicos e cadros de mando con R para a xestión de datos. Modelos de regresión para datos de alta dimensión. Modelos de clasificación para datos de alta dimensión.

Planificación				
Metodoloxías / probas	Competencias / Resultados	Horas lectivas (presenciais e virtuais)	Horas traballo autónomo	Horas totais
Prácticas a través de TIC	A1 A6 A7 A8 A13 A20 A21 A22	13	0	13
Estudo de casos	A1 A6 A7 A8 A13 A20 A21 A22	7	0	7
Traballos tutelados	A1 A6 A7 A8 A13 A20 A21 A22	0	101	101



Proba obxectiva	A1 A6 A7 A8 A13 A20 A21 A22	0	7	7
Sesión maxistral	A1 A6 A7 A8 A13 A20 A21 A22	21	0	21
Atención personalizada		1	0	1

\*Os datos que aparecen na táboa de planificación son de carácter orientativo, considerando a heteroxeneidade do alumnado

Metodoloxías	
Metodoloxías	Descrición
Prácticas a través de TIC	Metodoloxía que permite ao alumnado aprender de forma efectiva, a través de actividades de carácter práctico (demostracións, simulacións, etc.) a teoría dun ámbito de coñecemento, mediante a utilización das tecnoloxías da información e as comunicacións. As TIC supoñen un excelente soporte e canal para o tratamento da información e aplicación práctica de coñecementos, facilitando a aprendizaxe e o desenvolvemento de habilidades por parte do alumnado.
Estudo de casos	Metodoloxía onde o suxeito se enfrenta ante a descrición dunha situación específica que suscita un problema que ten que ser comprendido, valorado e resolto por un grupo de persoas, a través dun proceso de discusión. O alumno sitúase ante un problema concreto (caso), que lle describe unha situación real da vida profesional, e debe ser capaz de analizar unha serie de feitos, referentes a un campo particular do coñecemento ou da acción, para chegar a unha decisión razoada a través dun proceso de discusión en pequenos grupos de traballo.
Traballos tutelados	Metodoloxía deseñada para promover a aprendizaxe autónoma dos estudantes, baixo a tutela do profesor e en escenarios variados (académicos e profesionais). Está referida prioritariamente ao aprendizaxe do ?cómo facer as cousas?. Constitúe unha opción baseada na asunción polos estudantes da responsabilidade pola súa propia aprendizaxe. Este sistema de ensino baséase en dous elementos básicos: a aprendizaxe independente dos estudantes e o seguimento desa aprendizaxe polo profesor-titor.
Proba obxectiva	Proba escrita utilizada para a avaliación da aprendizaxe, cuxo trazo distintivo é a posibilidade de determinar se as respostas dadas son ou non correctas. Constitúe un instrumento de medida, elaborado rigorosamente, que permite avaliar coñecementos, capacidades, destrezas, rendemento, aptitudes, actitudes, intelixencia, etc. É de aplicación tanto para a avaliación diagnóstica, formativa como sumativa.  A Proba obxectiva pode combinar distintos tipos de preguntas: preguntas de resposta múltiple, de ordenación, de resposta breve, de discriminación, de completar e/ou de asociación. Tamén se pode construír con un só tipo dalgunha destas preguntas.
Sesión maxistral	Exposición oral complementada co uso de medios audiovisuais e a introdución de algunhas preguntas dirixidas aos estudantes, coa finalidade de transmitir coñecementos e facilitar a aprendizaxe. A clase maxistral é tamén coñecida como ?conferencia?, ?método expositivo? ou ?lección maxistral?. Esta última modalidade sóese reservar a un tipo especial de lección impartida por un profesor en ocasións especiais, cun contido que supón unha elaboración orixinal e baseada no uso case exclusivo da palabra como vía de transmisión da información á audiencia.

Atención personalizada	
Metodoloxías	Descrición
Traballos tutelados	Estímase que entre o alumnado haberá diferenzas notables tanto no que se refire á súa familiaridade cos conceptos e termos informáticos, como no que se refire ás habilidades para o manexo de ferramentas informáticas. Por este motivo, está previsto desenvolver unha atención personalizada para os traballos tutelados. A atención personalizada desenvolverase de forma individualizada durante as clases ou nas horas de tutoría.

Avaliación			
Metodoloxías	Competencias / Resultados	Descrición	Cualificación



Proba obxectiva	A1 A6 A7 A8 A13 A20 A21 A22	Proba individual sobre contido teórico-práctico	40
Traballos tutelados	A1 A6 A7 A8 A13 A20 A21 A22	Avaliarase a calidade dos traballos realizados	60

### Observacións avaliación

**PRIMEIRA OPORTUNIDADE** Para aprobar a materia é obrigatorio: UNHA NOTA MÍNIMA de 3 (sobre 6) en traballos tutelados. UNHA NOTA MÍNIMA de 2 (sobre 4) na proba obxectiva. Se non se obtén a nota mínima nos traballos tutelados ou na proba obxectiva, a nota máxima global da materia non será superior a 4,5. Calquera alumno que non se presente á proba obxectiva terá a cualificación de NON PRESENTADO.

**SEGUNDA OPORTUNIDADE** Só poderán optar á segunda oportunidade aqueles alumnos que non superen a materia na primeira oportunidade. As condicións son as mesmas que na primeira oportunidade, incluíndo a nota mínima en ambas as partes. Se un alumno/a decide non facer algunha das partes, manterá a cualificación obtida a primeira vez nesa parte. Calquera alumno que non opte pola recuperación de dúas partes algunha terá a cualificación de NON PRESENTADO.

**RENUNCIA ACADÉMICA** Dado que a asistencia ás clases presenciais non é obrigatoria, terán as mesmas condicións que o resto do alumnado aquel alumnado con matrícula a tempo parcial e con dispensa académica que o exima da asistencia ás clases.

**OPORTUNIDADE AVANZADA** Utilizaranse os mesmos criterios que nas oportunidades normais.

**IMPLICACIONES DO PLAXIO** A realización fraudulenta de probas ou actividades de avaliación, unha vez comprobada, implicará directamente unha cualificación de "0" na materia no momento oportuno.

### Fontes de información

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

### Recomendacións

#### Materias que se recomenda ter cursado previamente

Linguaxes de Programación, Análise e Consulta para a Xestión de Información/710G04029  
 Sistemas de Información para a Xestión Documental/710G04025  
 Informática para o Tratamento e Xestión da Información/710G04024

#### Materias que se recomenda cursar simultaneamente

#### Materias que continúan o temario

### Observacións

(\*A Guía docente é o documento onde se visualiza a proposta académica da UDC. Este documento é público e non se pode modificar, salvo casos excepcionais baixo a revisión do órgano competente dacordo coa normativa vixente que establece o proceso de elaboración de guías