



Teaching Guide				
Identifying Data				2022/23
Subject (*)	Big Data	Code	710G04035	
Study programme	Grao en Xestión Dixital de Información e Documentación			
Descriptors				
Cycle	Period	Year	Type	Credits
Graduate	1st four-month period	Fourth	Optional	6
Language	Spanish			
Teaching method	Face-to-face			
Prerequisites				
Department	Ciencias da Computación e Tecnoloxías da Información			
Coordinador	Cortiñas Álvarez, Alejandro	E-mail	alejandro.cortinas@udc.es	
Lecturers	Cortiñas Álvarez, Alejandro González Rueda, Ángel Manuel	E-mail	alejandro.cortinas@udc.es angel.manuel.rueda@udc.es	
Web				
General description	O obxectivo do curso é desenvolver coñecementos teóricos e prácticos sobre as tecnoloxías NoSQL e BigData, así como a análise e modelado estatístico.			

Study programme competences / results	
Code	Study programme competences / results
A1	CE1 - Know and understand the theoretical and methodological principles of information and documentation management to apply them in their professional activity
A6	CE6 - Search and retrieve information in various media to respond to the demand of information users
A7	CE7 - Plan and design an information management system, including information flows, both in an institutional and business context
A8	CE8 - Master the different methods of representation of data, information and knowledge that ensure efficient recovery
A13	CE13 - Know and master the techniques and regulations for the creation and authentication, meeting, selection, organization, representation, preservation, recovery, access, dissemination and exchange, and evaluation of resources and information services
A20	CE20 - Master the bases to develop research activities using multidisciplinary methods and principles
A21	CE21 - Possess knowledge of statistics and quantitative analysis of information
A22	CE22 - Acquire computational skills and management of new ICT

Learning outcomes			
Learning outcomes	Study programme competences / results		
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		
Capacidades 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		

Contents	
Topic	Sub-topic
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.

Planning				
Methodologies / tests	Competencies / Results	Teaching hours (in-person & virtual)	Student?s personal work hours	Total hours
ICT practicals	A1 A6 A7 A8 A13 A20 A21 A22	13	0	13
Case study	A1 A6 A7 A8 A13 A20 A21 A22	7	0	7
Supervised projects	A1 A6 A7 A8 A13 A20 A21 A22	0	101	101



Objective test	A1 A6 A7 A8 A13 A20 A21 A22	0	7	7
Guest lecture / keynote speech	A1 A6 A7 A8 A13 A20 A21 A22	21	0	21
Personalized attention		1	0	1
(*)The information in the planning table is for guidance only and does not take into account the heterogeneity of the students.				

Methodologies	
Methodologies	Description
ICT practicals	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.
Case study	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.
Supervised projects	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.
Objective test	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.
Guest lecture / keynote speech	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.

Personalized attention	
Methodologies	Description
Supervised projects	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.

Assessment			
Methodologies	Competencies / Results	Description	Qualification



Objective test	A1 A6 A7 A8 A13 A20 A21 A22	Proba individual sobre contido teórico-práctico	40
Supervised projects	A1 A6 A7 A8 A13 A20 A21 A22	Avaliarase a calidade dos traballos realizados	60

### Assessment comments

**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.

### Sources of information

<b>Basic</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>Complementary</b>	

### Recommendations

#### Subjects that it is recommended to have taken before

Programming, Analysis and Query Languages for Information Management/710G04029  
 Information Systems for Document Management/710G04025  
 Information Technology for the Treatment and Management of Information/710G04024

#### Subjects that are recommended to be taken simultaneously

#### Subjects that continue the syllabus

#### Other comments

(\* )The teaching guide is the document in which the URV publishes the information about all its courses. It is a public document and cannot be modified. Only in exceptional cases can it be revised by the competent agent or duly revised so that it is in line with current legislation.