



## Guía Docente

Datos Identificativos					2024/25
Asignatura (*)	Minaría de Datos	Código	710G04030		
Titulación	Grao en Xestión Dixital de Información e Documentación				
Descritores					
Ciclo	Período	Curso	Tipo	Créditos	
Grao	2º cuatrimestre	Terceiro	Optativa	6	
Idioma	CastelánGalego				
Modalidade docente	Presencial				
Prerrequisitos					
Departamento	Matemáticas				
Coordinación	Gómez Rodríguez, Marcos	Correo electrónico	marcos.gomez.rodriguez@udc.es		
Profesorado	Gómez Rodríguez, Marcos	Correo electrónico	marcos.gomez.rodriguez@udc.es		
Web					
Descrición xeral	O obxectivo fundamental desta materia é que o alumno coñeza os conceptos fundamentais e os principais modelos da minaría de datos, e a súa aplicación no campo das ciencias da información e documentación.				

## Competencias / Resultados do título

Código	Competencias / Resultados do título
--------	-------------------------------------

## Resultados da aprendizaxe

Resultados de aprendizaxe	Competencias / Resultados do título		
Adquisición de habilidades para a selección, tratamento, xestión e análise da información mediante técnicas de minaría de datos.			
Coñecemento e habilidades para a aplicación das principais técnicas de clasificación.			
Adquirir coñecementos de análises de datos computacional, incluíndo programas como o software estatístico R.			
Elixir as técnicas cuantitativas adecuadas aos obxectivos para tarefas de investigación, administración e xestión.			
Capacidade de análise e de síntese aplicada á xestión e organización da información.			
Coñecemento e habilidades para a aplicación de técnicas de regresión, detección de anomalías e series de tempo.			
Adquisición de habilidades para a toma de decisións a partir da análise estatística bases de datos complexas.			
Coñecemento e habilidades para a aplicación das principais técnicas de clasificación.			
Adquisición de habilidades para a toma de decisións a partir da análise estatística bases de datos complexas.			
Coñecemento e habilidades para a aplicación de técnicas de regresión, detección de anomalías e series de tempo.			
Capacidade de análise e de síntese aplicada á xestión e organización da información.			
Elixir as técnicas cuantitativas adecuadas aos obxectivos para tarefas de investigación, administración e xestión.			
Adquirir coñecementos de análises de datos computacional, incluíndo programas como o software estatístico R.			
Adquisición de habilidades para a selección, tratamento, xestión e análise da información mediante técnicas de minaría de datos.			
Adquisición de habilidades para a selección, tratamento, xestión e análise da información mediante técnicas de minaría de datos.			
Adquirir coñecementos de análises de datos computacional, incluíndo programas como o software estatístico R.			
Elixir as técnicas cuantitativas adecuadas aos obxectivos para tarefas de investigación, administración e xestión.			
Capacidade de análise e de síntese aplicada á xestión e organización da información.			
Coñecemento e habilidades para a aplicación de técnicas de regresión, detección de anomalías e series de tempo.			
Adquisición de habilidades para a toma de decisións a partir da análise estatística bases de datos complexas.			
Coñecemento e habilidades para a aplicación das principais técnicas de clasificación.			

## Contidos

Temas	Subtemas
-------	----------



Introducción á minería de datos.	<p>Conceptos preliminares.</p> <p>Tipos de problemas en minería de datos: descrición, clasificación, predición, clustering, detección de anomalías, etc.</p> <p>Tipos de aprendizaxe: supervisado e non supervisado.</p>
Métodos de clasificación non supervisada ou clúster	<p>Conceptos básicos.</p> <p>Métodos de clasificación xerárquica.</p> <p>Métodos de agrupamento por particións.</p> <p>Casos prácticos.</p>
Métodos de clasificación supervisada.	<p>Conceptos básicos.</p> <p>Modelos principais de clasificación supervisada ou recoñecemento de patróns.</p> <p>Validación de modelos de clasificación (como de ben predín?).</p> <p>Casos prácticos.</p>
Métodos avanzados de regresión.	<p>Introdución.</p> <p>Modelos de regresión univariantes e multivariantes.</p> <p>Selección de variables relevantes.</p> <p>Validación de modelos de regresión (como de ben se axusta aos datos?, como de ben fai predición?).</p> <p>Casos prácticos.</p>
Series de tempo	<p>Conceptos básicos.</p> <p>Análise descritivo de series de tempo.</p> <p>Uso práctico dos modelos de series de tempo.</p> <p>Casos prácticos.</p>
Técnicas estatísticas para minería de textos e recuperación da información.	<p>Conceptos básicos.</p> <p>Casos prácticos de aplicación da minería de textos.</p>

Planificación				
Metodoloxías / probas	Competencias / Resultados	Horas lectivas (presenciais e virtuais)	Horas traballo autónomo	Horas totais
Sesión maxistral		19	0	19
Prácticas a través de TIC		13	0	13
Estudo de casos		7	7	14
Traballos tutelados		1	101	102
Proba obxectiva		1	0	1
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
Sesión maxistral	Serán sesións expositivas nas que se introducirán e describirán os diversos temas da materia, mediante presentacións (usando os adecuados medios audiovisuais) que incluírán teoría e exemplos.
Prácticas a través de TIC	Desenvolveranse clases prácticas mediante software estatístico, nas que se introducirá a súa programación e aplicación a partir de casos reais e simulados.
Estudo de casos	Aplicaranse as técnicas estatísticas impartidas na materia para a resolución de exercicios e casos de estudo reais e simulados no ámbito da xestión dixital da información.
Traballos tutelados	Realizaranse traballos individuais e/ou en grupo, tutelados polos docentes da materia, nos que se abordará a resolución, mediante a aplicación de técnicas estatísticas e o software R, de exercicios prácticos ou de casos de estudo particulares relacionados co ámbito da comunicación e das ciencias da información. Tamén se poderá realizar un estudo de revisión acerca dun tema concreto da materia ou do software utilizado. Os traballos poderán ser propostos polos docentes ou polos propios alumnos (as propostas teranse en conta ou non sempre segundo o criterio do docente).



Proba obxectiva	Consistirá nunha proba tipo test sobre os contidos impartidos na materia.
-----------------	---

### Atención personalizada

Metodoloxías	Descrición
Prácticas a través de TIC Estudo de casos Sesión maxistral Traballos tutelados	Nas clases maxistrais se fomentará en todo momento o debate entre os alumnos e entre os alumnos e o profesor. Para a resolución de problemas será importante atender personalmente aos alumnos ante as posibles dúbidas que poidan xurdir. Esta atención servirá tamén, por unha parte, ao profesor para detectar posibles problemas na metodoloxía utilizada para impartir a asignatura e, por outra, aos alumnos para consolidar coñecementos teóricos e para expresar as súas inquedanzas acerca da asignatura. Será tamén fundamental a atención personalizada ao estudante durante as clases de prácticas TIC, sobre todo ata que non se familiarice co software estatístico a utilizar, ademais de na resolución de casos de estudo.

### Avaliación

Metodoloxías	Competencias / Resultados	Descrición	Cualificación
Prácticas a través de TIC		Valorarase a asistencia e/ou o desempeño dos alumnos co software estatístico.	20
Proba obxectiva		Proba tipo test consistente nun número de preguntas entre 10 e 20 con 3 respostas posibles.	40
Traballos tutelados		Realizaranse traballos individuais e/ou en grupo, tutelados polos docentes da materia, nos que se abordará a resolución, mediante a aplicación de técnicas estatísticas e o software R, de exercicios prácticos ou de casos de estudo particulares relacionados co ámbito da comunicación e das ciencias da información. Tamén se poderá realizar un estudo de revisión acerca dun tema concreto da materia ou do software utilizado. Os traballos poderán ser propostos polos docentes ou polos propios alumnos (as propostas serán tidas en conta ou non sempre segundo o criterio do docente).	40

### Observacións avaliación

**Primeira oportunidade**

Realizarase unha proba de resposta múltiple de 10 a 20 preguntas que representa o 40% da nota. Por outra banda, a avaliación continua constará da asistencia e/ou entrega de prácticas relacionadas coa aprendizaxe e aplicación do software estatístico e R para a resolución de problemas no campo da xestión dixital da información (20% da nota global), ademais da entrega dun ou varios traballos de aplicación da estatística para a resolución de casos de estudo en documentación dixital (alternativamente poderán ser traballos de revisión ou ampliación da materia) que representa o 40% da nota total.

**Segunda oportunidade**

Na avaliación da segunda oportunidade se seguirá o mesmo criterio que na primeira.

**Convocatoria adiantada**

Todas as observacións previas son aplicables aos estudantes que soliciten a convocatoria adiantada do exame.

**Calificación de non presentado**

En calquera das dúas oportunidades anuais figurará un NON PRESENTADO naqueles casos nos que o alumnado non acuda ó exame oficial da materia.

**Estudiante con recoñecemento de dedicación a tempo parcial e dispensa académica de exención de asistencia**

No caso do alumnado con recoñecemento de dedicación a tempo parcial e dispensa académica de exención de asistencia que decida non asistir a clases, este será avaliado nas dúas oportunidades como o resto do alumnado que se atopa nunha situación similar.

A realización fraudulenta das probas ou actividades de avaliación implicará directamente a calificación de suspenso "0" na materia na convocatoria correspondente, invalidando así calquera calificación obtida en todas as actividades de avaliación de cara á convocatoria extraordinaria.

### Fontes de información



<b>Bibliografía básica</b>	<ul style="list-style-type: none"> <li>- Silge, J. y Robinson, D. (2017). Text Mining with R: A Tidy Approach. O'Reilly</li> <li>- Cirillo, A. (2017). R Data Mining: Implement Data Mining Techniques Through Practical Use Cases and Real-world Datasets. . Packt Publishing.</li> <li>- Williams, G. (2011). Data mining with Rattle and R: The art of excavating data for knowledge discovery. . Springer Science &amp; Business Media.</li> <li>- Jockers, M.L. (2014). Text Analysis with R for Students of Literature. Springer</li> </ul>
<b>Bibliografía complementaria</b>	

## Recomendacións

### Materias que se recomenda ter cursado previamente

Ciencia de Datos/710G04026  
Fundamentos Estatísticos/710G04040

### Materias que se recomenda cursar simultaneamente

### Materias que continúan o temario

## Observacións

&lt;p&gt;Para axudar a conseguir unha contorna sostible e cumprir co obxectivo da acción número 5: ?Docencia e investigación saudable e sustentable&nbsp; ambiental e social? do &quot;Plan de Acción Green Campus Ferrol&quot;:&nbsp;&lt;p&gt;&lt;p&gt;1.- A entrega dos traballos documentais que se realicen nesta materia:&nbsp;&lt;p&gt;&lt;p&gt;1.1. Solicitarase en formato virtual e/ou soporte informático.&nbsp;&lt;p&gt;&lt;p&gt;1.2. Realizarase a través de Moodle, en formato dixital sen necesidade de imprimilos&nbsp;&lt;p&gt;&lt;p&gt;1.3. De se realizar en papel:&nbsp;&lt;p&gt;&lt;p&gt;- Non se empregarán plásticos.&nbsp;&lt;p&gt;&lt;p&gt;- Realizaranse impresións a dobre cara.&nbsp;&lt;p&gt;&lt;p&gt;- Empregarase papel reciclado.&nbsp;&lt;p&gt;&lt;p&gt;- Evitarase a impresión de borradores.&nbsp;&lt;p&gt;&lt;p&gt;2.- Débese facer un uso sostible dos recursos e a prevención de impactos negativos sobre o medio natural.&nbsp;&lt;p&gt;&lt;p&gt;3.- Débese ter en conta a importancia dos principios éticos relacionados cos valores da sostenibilidade nos comportamentos persoais e profesionais.&nbsp;&lt;p&gt;&lt;p&gt;4.- Segundo se recolle nas distintas normativas de aplicación para a docencia universitaria deberase incorporar a perspectiva de xénero nesta materia (usarase linguaxe non sexista, utilizarase bibliografía de autores de ambos os sexos, propiciarase a intervención en clase de alumnos e alumnas?).&lt;p&gt;&lt;p&gt;5.- Traballarase para identificar e modificar prexuízos e actitudes sexistas, e influirase na contorna para modificalos e fomentar valores de respecto e igualdade.&nbsp;&lt;p&gt;&lt;p&gt;6. Deberanse detectar situacións de discriminación por razón de xénero e proporanse accións e medidas para corrixilas.&nbsp;&lt;p&gt;&lt;p&gt;7. Facilitarase a plena integración do alumnado que por razón físicas, sensoriais, psíquicas ou socioculturais, experimenten dificultades a un acceso axeitado, igualitario e proveitoso á vida universitaria.&lt;p&gt;

(\*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