



## Guía Docente

Datos Identificativos					2023/24
Asignatura (*)	Ciencia de Datos	Código	710G04026		
Titulación					
Descritores					
Ciclo	Período	Curso	Tipo	Créditos	
Grao	2º cuatrimestre	Segundo	Formación básica	6	
Idioma	CastelánGalego				
Modalidade docente	Presencial				
Prerrequisitos					
Departamento	Matemáticas				
Coordinación	Francesch Domenech, María Érika	Correo electrónico	erika.francesch@udc.es		
Profesorado	Francesch Domenech, María Érika Tarrío Saavedra, Javier	Correo electrónico	erika.francesch@udc.es javier.tarrío@udc.es		
Web	<a href="https://estudios.udc.es/gl/study/start/710G04V01">https://estudios.udc.es/gl/study/start/710G04V01</a>				
Descrición xeral	Esta materia introduce e describe unha serie de conceptos estatísticos centrados na ciencia de datos. En concreto, comezase cunha introdución ás técnicas de mostraxe e deseño de enquisas, pasando pola inferencia estatística así como as principais técnicas de análises multivariante. Adicionalmente, introduciranse distintas ferramentas computacionais relacionadas co software estatístico R para a xeración de informes. Empregarase un enfoque principalmente aplicado, tratando de presentar todos os conceptos dentro do contexto da xestión da información e documentación.				

## Competencias / Resultados do título

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

## Resultados da aprendizaxe

Resultados de aprendizaxe	Competencias / Resultados do título		
Coñecemento das técnicas de inferencia básicas e adquisición de habilidades para a estimación e interpretación de intervalos de confianza e contraste de hipótese dunha e dúas poboacións.	A8 A13 A21	B1 B8 B9	
Coñecer os principais tipos de mostraxe e as ferramentas básicas para o deseño de enquisas.	A1 A13 A20 A21	B2 B3 B4 B5 B9	
Capacidade para comparar dúas ou máis poboacións a partir de bases de datos de diferente grao de complexidade.	A1 A21	B1 B2 B3 B4 B5	
Coñecemento das diferentes técnicas de análises de datos multivariantes para a descrición e obtención de información relevante a partir de bases de datos complexas.	A1 A20 A21	B1 B2 B3 B4 B5	
Habilidade para utilizar as ferramentas computacionais de análises de datos multivariantes.	A22	B11	C2 C6 C8



Integrar os coñecementos estatísticos teóricos e prácticos como vía para do coñecemento e pensamento reflexivo e totalizador.	A1 A13	B2 B3 B4 B5 B6 B7 B10	C4 C7 C8
Capacidade de análise e de síntese aplicada á xestión e organización da información.		B2 B3 B4 B5 B6 B7 B8 B9	C1 C3 C5
Adquisición de habilidades para a toma de decisións a partir da análise estatística bases de datos complexas.	A21	B2 B3 B8 B9	C8

Contidos	
Temas	Subtemas
Os temas seguintes desenvolven os contidos establecidos na ficha da Memoria de Verificación, sendo:	Introdución e conceptos estatísticos principais relativos a mostraxe e deseño de enquisas. Introdución á inferencia estatística e estimación puntual. Intervalos de confianza. Contraste de hipótese. Análise da varianza (ANOVA). Modelos de regresión. Outras técnicas de análises multivariante. Ferramentas computacionais para a xeración de informes estatísticos.
1. Mostraxe e enquisas: introdución e conceptos principais.	Conceptos xerais de mostraxe estatística e deseño de enquisas.
2. Introdución á inferencia estatística e estimación puntual.	Conceptos xerais. Mostraxe. Estimación de parámetros. Propiedades dos estimadores. Estimación puntual: estimación puntual da media, a varianza e dunha proporción.
3. Intervalos de confianza.	Concepto de Intervalo de confianza. Intervalo de confianza para unha media, para unha varianza, para unha proporción e para a diferenza de dúas medias.
4. Contraste de hipótese.	Conceptos xerais. Contrastes de hipóteses para a media, a proporción e para a diferenza de dúas medias. Contrastes de independencia.
5. Análise da varianza (ANOVA).	ANOVA gráfico. ANOVA dun factor. ANOVA de máis dun factor.
6. Modelos de regresión.	Modelo de regresión lineal simple e múltiple. Outros modelos de regresión.
7. Outras técnicas de análise multivariante: análise de compoñentes principais, análise factorial, análise de correspondencias, escalado multidimensional.	Introdución ás técnicas multivariantes máis usadas.
8. Ferramentas computacionais para a xeración de informes estatísticos.	Introdución a distintas ferramentas do software estatístico R para a xeración de informes: Rstudio, Rmarkdown, Gráficos con R, Htmlwidgets.

Planificación				
Metodoloxías / probas	Competencias / Resultados	Horas lectivas (presenciais e virtuais)	Horas traballo autónomo	Horas totais
Sesión maxistral	A1 A8 A20 A21 B1 B3 B7 C4 C7	21	0	21
Prácticas a través de TIC	A13 A22 B11 C2	12	0	12



Estudo de casos	A1 A8 A21 B2 B3 B4 B5 B6 B7 B8 B9 C1 C8	7	7	14
Traballos tutelados	B2 B4 B5 B6 B8 B9 B10 C1 C3 C5 C6 C8	1.02	100.98	102
Proba obxectiva	A21 B1 B2	1	0	1
Atención personalizada		0	0	0

\*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
Sesión maxistral Prácticas a través de TIC	A atención personalizada farase, globalmente, mediante titorías personalizadas directas e virtuais, individuais e grupais

Avaliación			
Metodoloxías	Competencias / Resultados	Descrición	Cualificación
Prácticas a través de TIC	A13 A22 B11 C2	Valorarase a asistencia e/ou o desempeño dos alumnos nas clases prácticas co software estatístico.	20
Proba obxectiva	A21 B1 B2	Proba tipo test consistente nun número de preguntas entre 10 e 20 con 3 respostas posibles.	40
Traballos tutelados	B2 B4 B5 B6 B8 B9 B10 C1 C3 C5 C6 C8	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 R para a resolución de problemas no campo da xestión dixital da información (20% da nota global), ademais da entrega dun e/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, leste 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>- Daniel Peña (2002). Análisis de datos multivariantes. S.A. MCGRAW-HILL / INTERAMERICANA DE ESPAÑA</li> <li>- Cao, R., Francisco, M., Naya, S., Presedo, M.A., Vázquez, M., Vilar, J.A. y Vilar, J.M. (2001). Introducción a la Estadística y sus aplicaciones. Pirámide</li> <li>- Egghe, L. y Rousseau, R. (1990). Introduction to Infometrics. Quantitative Methods in Library, Documentation and Information Science. Amsterdam: Elsevier</li> <li>- Everitt, B. y Hothorn, T. (2011). An Introduction to Applied Multivariate Analysis with R. Springer-Verlag New York</li> </ul>
<b>Bibliografía complementaria</b>	<ul style="list-style-type: none"> <li>- Daniel Zelterman (2015). Applied Multivariate Statistics with R. Springer International Publishing</li> <li>- Cástor Guisande, Antonio Vaamonde (2012). Gráficos estadísticos y mapas con R. Díaz de Santos</li> <li>- Vélez, R. &amp; García, A. (1993). Principios de Inferencia Estadística. UNED</li> </ul>

## Recomendacións

### Materias que se recomenda ter cursado previamente

Fundamentos Estatísticos/710G04040

### Materias que se recomenda cursar simultaneamente

### Materias que continúan o temario

Minaría de Datos/710G04030

## Observacións

Para axudar a conseguir unha contorna inmediata sostida e cumprir co obxectivo da acción número 5: ?Docencia e investigación saudable e sustentable ambiental e social? do "Plan de Acción Green Campus Ferrol:A entrega dos traballos documentais que se realicen nesta materia: ? Solicitaranse en formato virtual e/ou soporte informático. ? Realizarase a través de Moodle, en formato dixital sen necesidade de imprimilos. ? En caso de ser necesario realízalos en papel:- Non se empregarán plásticos.- Realizaranse impresións a dobre cara.- Empregarase papel reciclado.- Evitarase a impresión de borradores. ? Débese de facer un uso sustentable&nbsp; dos recursos e a prevención de impactos negativos sobre o medio natural. ? Traballarase para identificar e modificar prexuízos e actitudes sexistas, e influirase na contorna para modificalos e fomentar valores de respecto e igualdade. ? Deberanse detectar situacións de discriminación e propoñeranse accións e medidas para corrixilas.

(\* )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

