



Teaching Guide				
Identifying Data				2020/21
Subject (*)	Data Science	Code	710G04026	
Study programme	Grao en Xestión Dixital de Información e Documentación			
Descriptors				
Cycle	Period	Year	Type	Credits
Graduate	2nd four-month period	Second	Basic training	6
Language	Spanish			
Teaching method	Hybrid			
Prerequisites				
Department	Matemáticas			
Coordinador	González Rueda, Ángel Manuel	E-mail	angel.manuel.rueda@udc.es	
Lecturers	González Rueda, Ángel Manuel	E-mail	angel.manuel.rueda@udc.es	
Web	estudios.udc.es/gl/subject/710G04V01/710G04026/2020			
General description	<p>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.</p>			
Contingency plan	<ol style="list-style-type: none"> <li>1. Modifications to the contents</li> <li>2. Methodologies <ul style="list-style-type: none"> <li>*Teaching methodologies that are maintained</li> <li>*Teaching methodologies that are modified</li> </ul> </li> <li>3. Mechanisms for personalized attention to students</li> <li>4. Modifications in the evaluation <ul style="list-style-type: none"> <li>*Evaluation observations:</li> </ul> </li> <li>5. Modifications to the bibliography or webgraphy</li> </ol>			

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
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
B1	CB1 - Possess and understand knowledge that provides a basis or opportunity to be original in the development and / or application of ideas, often in a research context
B2	CB2 - Apply the knowledge acquired and their ability to solve problems in new or unfamiliar environments within broader (or multidisciplinary) contexts related to their area of study



B3	CB3 - Be able to integrate knowledge and face the complexity of making judgments based on information that, being incomplete or limited, includes reflections on social and ethical responsibilities linked to the application of their knowledge and judgments
B4	CB4 - Know how to communicate their conclusions -and the knowledge and ultimate reasons that sustain them- to specialized and non-specialized audiences in a clear and unambiguous way
B5	CB5 - Possess the learning skills that allow them to continue studying in a way that will be largely self-directed or autonomous
B6	CG1 - Capacity for cooperation, teamwork and collaborative learning
B7	CG2 - Capacity for reflection and critical reasoning
B8	CG3 - Capacity for planning, organization and management of resources, information and operations
B9	CG4 - Capacity for analysis, diagnosis and decision making
B10	CG5 - Ability to work in an international and global context
B11	CG6 - Ability to understand the importance, value and function of the Digital Information and Documentation Management in the current ICT society
C1	CT1 - Express correctly, both orally and in writing, in the official languages ??of the autonomous community
C2	CT2 - Use the basic tools of information and communication technologies (ICT) necessary for the exercise of their profession and for learning throughout their lives
C3	CT3 - Develop oneself for the exercise of a citizenship that respects democratic culture, human rights and the gender perspective
C4	CT4 - Understand the importance of the entrepreneurial culture and know the means available to entrepreneurs
C5	CT5 - Acquire skills for life and habits, routines and healthy lifestyles
C6	CT6 - Develop the ability to work in interdisciplinary or transdisciplinary teams, to offer proposals that contribute to a sustainable environmental, economic, political and social development
C7	CT7 - Assess the importance of research, innovation and technological development in the socio-economic and cultural progress of society
C8	CT8 - Have the ability to manage time and resources: develop plans, prioritize activities, identify criticisms, establish deadlines and comply with them

Learning outcomes			
Learning outcomes	Study programme competences / results		
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

Contents	
Topic	Sub-topic
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.

Planning				
Methodologies / tests	Competencies / Results	Teaching hours (in-person & virtual)	Student?s personal work hours	Total hours
Guest lecture / keynote speech	A1 A8 A20 A21 B1 B3 B7 C7 C4	21	0	21
ICT practicals	A13 A22 B11 C2	12	0	12



Case study	A1 A8 A21 B2 B3 B4 B5 B6 B7 B8 B9 C1 C8	7	7	14
Supervised projects	B2 B4 B5 B6 B8 B9 B10 C1 C3 C5 C6 C8	1.02	100.98	102
Objective test	A21 B1 B2	1	0	1
Personalized attention		0	0	0

(\*)The information in the planning table is for guidance only and does not take into account the heterogeneity of the students.

Methodologies	
Methodologies	Description
Guest lecture / keynote speech	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.
ICT practicals	Desenvolveranse clases prácticas mediante software estatístico, nas que se introducirá a súa programación e aplicación a partir de casos reais e simulados.
Case study	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.
Supervised projects	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).
Objective test	Consistirá nunha proba tipo test sobre os contidos impartidos na materia.

Personalized attention	
Methodologies	Description
Guest lecture / keynote speech ICT practicals	A atención personalizada farase, globalmente, mediante titorías personalizadas directas e virtuais, individuais e grupais

Assessment			
Methodologies	Competencies / Results	Description	Qualification
ICT practicals	A13 A22 B11 C2	Valorarase a asistencia e/ou o desempeño dos alumnos nas clases prácticas co software estatístico.	20
Objective test	A21 B1 B2	Proba tipo test consistente nun número de preguntas entre 10 e 20 con 3 respostas posibles.	40
Supervised projects	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

Assessment comments



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. 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. Segunda oportunidade Na avaliación da segunda oportunidade seguirase o mesmo criterio que na primeira.

### Sources of information

<b>Basic</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>Complementary</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>

### Recommendations

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

Fundamentals of Statistics/710G04040

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

#### Subjects that continue the syllabus

### Other comments

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 prexuizos 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.

(\*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.