UNIVERSIDADE DA CORUÑA

Guia docente					
Datos Identificativos					2022/23
Asignatura (*)	Análisis Estadístico de Datos con Dependencia			Código	614G02022
Titulación	Grao en Ciencia e Enxeñaría de Datos				
Descriptores					
Ciclo	Periodo	Curso		Tipo	Créditos
Grado	1º cuatrimestre	Tercero		Obligatoria	6
Idioma	Castellano				
Modalidad docente	Presencial				
Prerrequisitos					
Departamento	Matemáticas				
Coordinador/a	Aneiros Perez, German		Correo electrónico	german.aneiros@udc.es	
Profesorado	Aneiros Perez, German Fernández Casal, Rubén		Correo electrónico	german.aneiros@udc.es ruben.fcasal@udc.es	
Web					
Descripción general	Modelización de datos que presentan dependencia temporal y/o espacial. Se utilizarán modelos estadísticos paramétricos. A partir de dichos modelos, se obtendrán predicciones de valores en nuevos instantes temporales o posiciones espaciales. Se hará uso del software R.				

Competencias / Resultados del título	
Código	Competencias / Resultados del título
A17	CE17 - Capacidad para la construcción, validación y aplicación de un modelo estocástico de un sistema real a partir de los datos observados y el análisis crítico de los resultados obtenidos.
A19	CE19 - Capacidad para comprender, plantear, formular y resolver problemas susceptibles de ser abordados a través de modelos estadísticos para datos que presentan dependencia.
A20	CE20 - Conocimiento de las herramientas informáticas en el campo del análisis de los datos y modelización estadística, y capacidad para seleccionar las más adecuadas para la resolución de problemas.
B2	CB2 - Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio
B3	CB3 - Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética
B4	CB4 - Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado
B7	CG2 - Elaborar adecuadamente y con cierta originalidad composiciones escritas o argumentos motivados, redactar planes, proyectos de trabajo, artículos científicos y formular hipótesis razonables.
B8	CG3 - Ser capaz de mantener y extender planteamientos teóricos fundados para permitir la introducción y explotación de tecnologías nuevas y avanzadas en el campo.
B9	CG4 - Capacidad para abordar con éxito todas las etapas de un proyecto de análisis de datos: exploración previa de los datos, preprocesado, análisis, visualización y comunicación de resultados.
B10	CG5 - Ser capaz de trabajar en equipo, especialmente de carácter multidisciplinar, y ser hábiles en la gestión del tiempo, personas y toma de decisiones.
C1	CT1 - Utilizar las herramientas básicas de las tecnologías de la información y las comunicaciones (TIC) necesarias para el ejercicio de su profesión y para el aprendizaje a lo largo de su vida.
C4	CT4 - Valorar la importancia que tiene la investigación, la innovación y el desarrollo tecnológico en el avance socioeconómico y cultural de la sociedad.

Resultados de aprendizaje	
Resultados de aprendizaje	Competencias / Resultados del título

Conocer y comprender conceptos generales relativos a procesos estocásticos.	A17 A19	B4 B8	
Identificar y analizar modelos estadísticos susceptibles de haber generado a un conjunto de datos dependientes.	A17 A19	B2 B4 B9 B10	C4
Conocer y aplicar técnicas de estimación de los parámetros presentes en los modelos estadísticos con datos dependientes.	A17 A19	B2 B4 B9 B10	C4
Entender la importancia de llevar a cabo una diagnosis de un modelo construido con datos que presenten dependencia temporal y/o espacial.	A17 A19	B4	C4
Conocer y aplicar los fundamentos de la predicción en series de tiempo.	A17 A19	B2 B4 B7 B9 B10	C4
Ser capaz de interpretar los modelos propuestos y los resultados obtenidos al utilizar técnicas estadísticas para datos dependientes.		B3 B4 B7 B9 B10	C4
Saber manejar con soltura programas informáticos avanzados de análisis estadístico.	A20		C1

Contenidos	
Tema	**Subtema**
1. Análisis descriptivo de una serie de tiempo	1.1 Introducción 1.2 Descomposición de una serie de tiempo
2. Series de tiempo y procesos estocásticos	2.1 Introducción 2.2 Procesos estocásticos: concepto y definiciones asociadas
3. Metodología Box-Jenkins	3.1 Introducción 3.2 Modelización ARIMA y predicción
4. Tópicos adicionales	Introducción a análisis de intervención, valores atípicos, regresión con series de tiempo, y cluster y clasificación con series de tiempo
5. Estadística Espacial	5.1 Tipos de procesos espaciales 5.2 Análisis exploratorio de datos espaciales 5.3 Modelado de procesos geoestadísticos 5.4 Predicción Kriging 5.5 Introducción a la Estadística espacio-temporal

Planificación				
Metodologías / pruebas	**Competencias / Resultados**	**Horas lectivas (presenciales y virtuales)**	**Horas trabajo autónomo**	**Horas totales**
Sesión magistral	A17 A19 B3 B4 B8 B9 C4	21	42	63
Prácticas a través de TIC	A20 B2 B4 B7 B9 B10 C1 C4	14	42	56
Seminario	B7 B8 C4	7	14	21

Prueba mixta	A17 A19 B2 B3 B8 B9 C4	2	0	2
Atención personalizada		8	0	8
(*)Los datos que aparecen en la tabla de planificación són de carácter orientativo, considerando la heterogeneidad de los alumnos				

Metodologías	
Metodologías	**Descripción**
Sesión magistral	El profesor, con la ayuda de los medios audiovisuales pertinentes, expondrá los contenidos teórico-prácticos de la asignatura. El material empleado estará a disposición de los estudiantes.
Prácticas a través de TIC	Se trata de poner en práctica la metodología estadística aprendida en las sesiones magistrales. Se hará uso del software R (gratuito). Se le harán llegar al estudiante distintos conjuntos de datos reales para ser analizados a lo largo de las sesiones de prácticas. A su vez, se le recomendará que obtenga otros datos por su cuenta para que refuerce su destreza en el análisis de datos con dependencia, esta vez en las horas dedicadas al trabajo autónomo.
Seminario	Su objetivo es apuntalar, a través de la realización de ejercicios teórico-prácticos, la comprensión de los contenidos expuestos en las sesiones magistrales. En algunas ocasiones, podrían dedicarse a completar las prácticas a través de TIC. Se le harán llegar al estudiante distintos ejercicios para que trate de resolverlos (algunos de ellos a lo largo del propio seminario y otros en las horas dedicadas al trabajo autónomo).
Prueba mixta	Prueba tipo test a través de la cual se evaluará el grado de aprendizaje alcanzado por el estudiante.

Atención personalizada	
Metodologías	**Descripción**
Sesión magistral Prácticas a través de TIC Seminario	A lo largo de cualquiera de las actividades en las que el profesor esté presente (sesión magistral, prácticas a través de TIC, seminario), el estudiante tiene la oportunidad de exponerle las dudas que le surjan acerca de los contenidos que se están impartiendo o llevando a la práctica. Si la resolución de dichas dudas implica una clara pérdida de continuidad de la actividad que se está llevando a cabo, o bien si las dudas le surgen al estudiante en horas no presenciales, puede hacer uso de las tutorías individualizadas para que le sean resueltas por parte del profesor.

Evaluación			
Metodologías	**Competencias / Resultados**	**Descripción**	**Calificación**
Prueba mixta	A17 A19 B2 B3 B8 B9 C4	Consistirá en una prueba tipo test de conocimientos teórico-prácticos, que será realizada en la fecha oficial aprobada por la Junta de Facultad.	60
Prácticas a través de TIC	A20 B2 B4 B7 B9 B10 C1 C4	Parte del grado de aprendizaje alcanzado en las prácticas a través de TIC se evaluará de manera continua. Para ello, se realizarán una o dos pruebas a lo largo del curso (en el horario establecido para dichas prácticas) y/o uno o dos trabajos (individuales o en grupo). Su resolución requerirá del uso del software R y también del conocimiento de la metodología estadística explicada en las sesiones magistrales.	40

Observaciones evaluación

Primera oportunidad: Si se denotan a través de P y F las notas (sobre 10) obtenidas en las "Prácticas a través de TIC" y en la "Prueba mixta", respectivamente, la nota final será:

(a) 0.4*P+0.6*F, siempre y cuando tanto P como F sean mayores o iguales que 3. (b) 4.5*(0.4*P+0.6*F)/7.2, siempre y cuando P y/o F sean menores que 3.Como consecuencia de lo anterior, se tiene que para superar la asignatura será necesario obtener un mínimo de 3 puntos tanto en "Prácticas a través de TIC" como en la "Prueba mixta".Se avisará de la fecha en que se realizará cada prueba práctica con una antelación mínima de 7 días. Los trabajos prácticos tendrán un plazo de entrega de al menos 7 días.Segunda oportunidad: La puntuación obtenida en "Prácticas a través de TIC" en la primera oportunidad se mantendrá para esta segunda oportunidad si es superior o igual a 5. Si es menor que 5, el alumno tendrá la opción de volver a hacer la(s) prueba(s)/trabajo(s) correspondiente(s) a la(s) parte(s) de prácticas no superada(s); si no hace uso de dicha opción, se le mantendrá la nota de la primera oportunidad. Todos los alumnos deben realizar la "Prueba mixta". La nota final se obtendrá de la misma forma que en la primera oportunidad.

Fuentes de información	
Básica	- Bivand R.S., Pebesma E.J., Gómez-Rubio V. (2008). Applied Spatial Data Analysis with R. Springer - Cowpertwait P.S.P., Metcalfe A.V. (2009). Introductory Time Series with R. Springer - Cryer J.D., Chan K-S. (2008). Time Series Analysis. With Applications in R. Springer (2ª edición) - Diggle P., Ribeiro P.J. (2007). Model-based Geostatistics. Springer - Hyndman R.J., Athanasopoulos G. (2018). Forecasting: Principles and Practice. O Texts (2ª edición) (accesible online en https://otexts.com/fpp2/) - Montero P., Vilar J.A. (2014). TSclust: An R package for time series clustering. Journal of Statistical Software 62(1)
Complementária	- Brockwell P.J., Davis R.A. (2002). Introduction to Time Series and Forecasting. Springer (2ª edición) - Chilès J.P., Delfiner P. (1999). Geostatistics: modeling spatial uncertainty. Wiley - Liao T.W. (2005). Clustering of time series-a survey. Pattern Recognittion 38, 1857-1874 - Peña D. (2005). Análisis de Series Temporales. Alianza Editorial - Shumway R.H., Stoffer D.S. (2017). Time Series Analysis and Its Applications. With R Examples. Springer (4ª edición) - Wikle C.K., Zammit-Mangion A., Cressie N. (2019). Spatio-temporal Statistics with R. Chapman and Hall/CRC (accesible online en https://spacetimewithr.org)

Recomendaciones
Asignaturas que se recomienda haber cursado previamente
Modelos de Regresión/614G02012 Modelización Estadística de Datos de Alta Dimensión/614G02013 Inferencia Estadística/614G02007 Probabilidad y Estadística Básica/614G02003
Asignaturas que se recomienda cursar simultáneamente
Asignaturas que continúan el temario
Técnicas de Simulación y Remuestreo/614G02036 Análisis Estadístico de Datos Complejos/614G02031
Otros comentarios

(*) La Guía Docente es el documento donde se visualiza la propuesta académica de la UDC. Este documento es público y no se puede modificar, salvo cosas excepcionales bajo la revisión del órgano competente de acuerdo a la normativa vigente que establece el proceso de elaboración de guías