



DATOS IDENTIFICATIVOS

Análise Exploratoria de Datos

Materia	Análise Exploratoria de Datos			
Código	V03M100V01104			
Titulación	Máster Universitario en Técnicas Estadísticas			
Descritores	Creditos ECTS	Sinale	Curso	Cuadrimestre
	6	OB	1	1c
Lingua de impartición	Castelán Galego			
Departamento	Estatística e investigación operativa			
Coordinador/a	Roca Pardiñas, Javier			
Profesorado	Iglesias Perez, Maria Carmen Perez Gonzalez, Ana Roca Pardiñas, Javier			
Correo-e	roca@uvigo.es			
Web	http://eio.usc.es/pub/mte/			
Descrición xeral	Profesorado: Javier Roca Pardiñas (UVigo): 1.5 ECTS Ana Pérez González (UVigo): 3 ECTS M. Carmen Iglesias Pérez (UVigo): 1.5 ECTS			
	Más información en http://eio.usc.es/pub/mte/			

Competencias de titulación

Código	
A1	Capacidad para comprender, presentar, formular y resolver aquellos problemas susceptibles de ser abordados a través de modelos de la estadística
A3	Conocer las aplicaciones de los modelos de la estadística
A5	Coñecer algoritmos de resolución de los problemas y manejar el software adecuado

Competencias de materia

Resultados previstos na materia	Tipoloxía	Resultados de Formación e Aprendizaxe
1. Comprender la finalidad, resultados y beneficios del análisis de un conjunto de datos, así como de sus requerimientos, con el fin de permitir una mejor modelización de problemas y experimentos.	saber saber hacer	A1 A3
2. Adquirir una visión general de los grandes grupos de problemas en la minería de datos y de los grupos de técnicas más apropiadas para resolverlos.	saber	A1 A3 A5
3. Conocer la metodología genérica de un proyecto de minería de datos en el que se encuadran las diferentes técnicas estadísticas que se tratan en el máster.	saber	A1
4. Conocer los conceptos, y técnicas numéricas y gráficas, esenciales para el descubrimiento y entendimiento de las estructuras y relaciones contenidas en un conjunto de datos.	saber	A1 A3 A5
5. Descubrir la problemática del análisis de un conjunto de datos, con el fin de motivar conceptos y técnicas de la teoría de la probabilidad y de la estadística matemática.	saber	A1 A3 A5
6. Conocer el panorama de software disponible para la minería de datos, así como saber aplicar cada técnica estudiada mediante alguna herramienta informática.	saber saber hacer	A1 A3 A5

Contidos

Tema

1. INTRODUCCIÓN AL ANÁLISIS EXPLORATORIO DE DATOS:	CONCEPTOS DE POBLACIÓN, MUESTRA, TIPOS DE VARIABLES. MANEJO DEL SOFTWARE ESTADÍSTICO R.
2. ANÁLISIS DESCRIPTIVO UNIDIMENSIONAL:	MEDIDAS DESCRIPTIVAS NUMÉRICAS, CREACIÓN DE GRÁFICOS, DETECCIÓN DE DATOS ATÍPICOS, ESTUDIO DE LA NORMALIDAD DE LOS DATOS.
3. ANÁLISIS DE DOS O MÁS VARIABLES:	TABLAS DE CORRELACIÓN Y DE CONTINGENCIA. GRÁFICOS PARA DOS VARIABLES. ANÁLISIS DE VARIABLES CUALITATIVAS. MEDIDAS DE ASOCIACIÓN Y DE CORRELACIÓN. COMPARACIÓN DE DISTRIBUCIONES: TEST PARAMÉTRICOS Y NO PARAMÉTRICOS.
4. ESTUDIO DE LAS DISTRIBUCIONES CONDICIONADAS;	ANÁLISIS DE LA VARIANZA, MODELO DE REGRESIÓN LINEAL SIMPLE Y MÚLTIPLE, MODELO DE REGRESIÓN PARAMÉTRICOS NO LINEALES, REGRESIÓN LOGÍSTICA. MODELOS DE REGRESIÓN NO PARAMÉTRICOS.
5. INTRODUCCIÓN AL ANÁLISIS DE DATOS MULTIVARIANTES	ANÁLISIS DE COMPONENTES PRINCIPALES, ANÁLISIS FACTORIAL, ANÁLISIS DE CONGLOMERADOS Y ANÁLISIS DISCRIMINANTE.
6. ESTIMACIÓN EN SITUACIONES ESPECIALES:	DATOS FALTANTES, DATOS ATÍPICOS E DATOS CENSURADOS.
7. INTRODUCCIÓN AL ANÁLISIS EXPLORATORIO DE SERIES TEMPORALES Y DEPENDENCIA ESPACIAL. DATOS DEPENDIENTES:	

Planificación

	Horas na aula	Horas fóra da aula	Horas totais
Prácticas de laboratorio	12	30	42
Titoría en grupo	0	20	20
Sesión maxistral	22	38	60
Resolución de problemas e/ou exercicios	10	10	20
Probas prácticas, de execución de tarefas reais e/ou simuladas.	4	4	8

*Os datos que aparecen na táboa de planificación son de carácter orientador, considerando a heteroxeneidade do alumnado.

Metodoloxía docente

	Descrición
Prácticas de laboratorio	La docencia se desenvolverá mediante a resolución de problemas reais ou simulados utilizando os modelos tratados en las sesiones magistrales. Se utilizará principalmente software R.
Titoría en grupo	Se mantendrá un servizo de tutoría en grupo a los alumnos.
Sesión maxistral	La docencia se desenvolverá mediante a exposición por parte del profesor de las diferentes técnicas de Análisis Exploratorio de Datos Para ello, los alumnos dispondrán de apuntes elaborados que servirán de material básico para el estudio y en su defecto de material e información sobre bibliografía específica disponible en la biblioteca o en internet.

Atención personalizada

Metodoloxías Descrición

Titoría en grupo	La tutoría se desenvolverá intercalada entre las sesiones magistrales y las prácticas de laboratorio con el fin adicional de evaluar el grado de asimilación de la materia por parte de los alumnos, así como mediante correo electrónico individual o mediante un foro en el que se publicitarán las preguntas formuladas y las respuestas proporcionadas por el profesor.
------------------	---

Avaliación

	Descrición	Cualificación
Resolución de problemas e/ou exercicios	Realización de cuestionarios e actividades que sexan plantexadas durante OI periodo de docencia da materia.	40
Probas prácticas, de execución de tarefas reais e/ou simuladas.	Realización dun exame final teórico práctico.	60

Outros comentarios sobre a Avaliación

Bibliografía. Fontes de información

BIBLIOGRAFÍA BÁSICA

· Abraira Santos, V. y Pérez de Vargas, A. (1996). Métodos Multivariantes en Bioestadística. Centro de Estudios Ramón Areces.
· Crawley, M. J. (2005) Statistics: an introduction using R. Ed. John Wiley and Sons.

- Everitt, B. S. (2005) An R and S-PLUS Companion to Multivariate Analysis. Ed. Springer
- Guisande González, C. y Vaamonde Liste A. (2011). Tratamiento de datos con R, Statistica y SPSS. Ed. Diaz de Santos Editores.
- Hair J. F., Anderson, Tatham and Black (1996). Análisis multivariante. 5ª Edición. Prentice Hall.
- Martínez Almécija, Alfredo y otros (1993). Inferencia Estadística. Un enfoque clásico. Pirámide.
- Peña Sánchez de Rivera, D. (1994). Estadística. Modelos y Métodos. 1. Fundamentos. Alianza Universidad Textos.
- Peña Sánchez de Rivera, D. (1999). Estadística. Modelos y Métodos. 2. Modelos Lineales y Series Temporales. Alianza Universidad Textos.
- TUKEY, J.W. (1977). Exploratory Data Analysis. Addison-Wesley
- Zuur, Alain F. (2009) A Beginner's guide to R. New York . Springer.

BIBLIOGRAFÍA COMPLEMENTARIA

- AGRESTI, A. (1996): An Introduction to Categorical Data Analysis, John Wiley, Nueva York.
- Cowpertwait, P.S.P. and Metcalfe, A. V. (2009). Introductory Time Series with R. Ed. Springer.
- Maindonald, J. H. (2007) Data analysis and graphics using R : an example-based approach. Cambridge University Press.
- Rousseeuw P.J., Leroy A.M. (2003) Robust regression and outlier detection John Wiley and Sons.
- Sheather, S. J . (2009) A modern approach to regression with R. New York ; London : Springer.

Recomendaciones

Outros comentarios

No se necesita haber cursado ninguna otra asignatura del máster. Una de las recomendaciones fundamentales para la superación de esta materia es la asistencia regular a las clases. Es muy importante el seguimiento del trabajo realizado en el aula.

Los requisitos básicos de esta materia sólo requieren un uso a nivel usuario de Windows y un conocimiento básico de Estadística.

El software utilizado será el software libre R.

En algunas sesiones se realizarán cuestionarios y ejercicios sobre temas impartidos para la valoración de la evolución y comprensión de los alumnos sobre la materia.
