



## DATOS IDENTIFICATIVOS

### Análise exploratorio de datos

Materia	Análise exploratorio de datos			
Código	V03M184V01101			
Titulación	Máster Universitario en Técnicas Estatísticas			
Descritores	Creditos ECTS	Sinale	Curso	Cuadrimestre
	5	OB	1	1c
Lingua de impartición				
Departamento				
Coordinador/a	Roca Pardiñas, Javier			
Profesorado	Roca Pardiñas, Javier			
Correo-e	roca@uvigo.es			
Web	<a href="http://eio.usc.es/pub/mte/">http://eio.usc.es/pub/mte/</a>			
Descrición xeral				

## Competencias

Código	
C1	Coñecer, identificar, modelar, estudar e resolver problemas complexos de estatística e investigación operativa, nun contexto científico, tecnolóxico ou profesional, xurdidos en aplicacións reais.
C2	Desenvolver autonomía para a resolución práctica de problemas complexos xurdidos en aplicacións reais e para a interpretación dos resultados de face á axuda en tómaa de decisións.
C6	Adquirir coñecementos teóricos e prácticos avanzados de diferentes técnicas matemáticas, dirixidas especificamente para a toma de decisións, e desenvolver a capacidade de reflexión para avaliar e decidir entre diferentes perspectivas en contextos complexos.
C8	Adquirir coñecementos teórico-prácticos avanzados de técnicas dirixidas a facer inferencias e contrastes con variables e parámetros dun modelo estatístico e saber aplicarlles con autonomía suficiente nun contexto científico, tecnolóxico ou profesional.
C9	Coñecer e saber aplicar de xeito autónomo en contextos científicos, tecnolóxicos ou profesionais, técnicas de aprendizaxe automática e técnicas de análise de datos en alta dimensión (big data).
C10	Adquirir coñecementos avanzados sobre metodoloxías para a obtención e procesamento de datos de diferentes fontes, como enquisas, Internet ou ""na nube"".

## Resultados de aprendizaxe

Resultados previstos na materia	Resultados de Formación e Aprendizaxe
Comprender a finalidade, resultados e beneficios da análise dun conxunto de datos, así como dos seus requirimentos, co fin de permitir unha mellor modelización de problemas e experimentos.	C1 C2
Descubrir a problemática da análise dun conxunto de datos, co fin de motivar conceptos e técnicas da teoría da probabilidade e da estatística matemática.	C6 C8 C9 C10
Coñecer os conceptos, e técnicas numéricas e gráficas, esenciais para o descubrimento e entendemento das estruturas e relacións contidas nun conxunto de datos	C1 C2 C6
Adquirir unha visión xeral dalgúns dos principais problemas estatísticos e dos grupos de técnicas máis apropiadas para resolvelos.	C1 C2 C6

## Contidos

Tema	
Manexo do software estatístico R.	Descarga e instalación. Lectura e tratamento de datos. Procedementos gráficos. Introducción á programación.
Natureza e tipo de datos.	Poboación e mostra. Tipos de mostraxe. Variables aleatorias discretas e continuas.
Medidas descritivas numéricas.	Media, desviación típica, varianza, mediana, rango, e cuantiles.
Gráficos estatísticos.	Principais gráficos estatísticos: gráfico de puntos, gráfico de barras, gráfico de sectores, histogramas e gráficos de densidade.
Tablas de frecuencias.	Construcción de tablas de frecuencias para uno y varios factores. Frecuencias absolutas, relativas y marginales. Construcción de tablas con R.
Estudo de correlación.	Definición de covarianza e correlación. Interpretación e representación gráfica. Tipos de correlación. Matrices e gráficos de correlación *multivariantes.

## Planificación

	Horas na aula	Horas fóra da aula	Horas totais
Seminario	0	12	12
Prácticas con apoio das TIC	15	15	30
Lección maxistral	25	23	48
Resolución de problemas e/ou exercicios	10	10	20
Práctica de laboratorio	0	15	15

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

## Metodoloxía docente

	Descrición
Seminario	Manterase un servizo de titoría en grupo aos alumnos. Os alumnos tamén poderán consultar as súas dúbidas por correo electrónico
Prácticas con apoio das TIC	A docencia desenvolverase mediante a resolución de problemas reais ou simulados utilizando os modelos tratados nas sesións maxistrais. Utilizarase principalmente o software R.
Lección maxistral	A docencia desenvolverase mediante a exposición por parte do profesor das diferentes técnicas de Análise Exploratorio de Datos Para iso, os alumnos disporán de apuntamentos elaborados que servirán de material básico para o estudo e na súa falta de material e información sobre bibliografía específica dispoñible na biblioteca ou en internet.

## Atención personalizada

### Metodoloxías Descrición

Seminario	As dúbidas dos alumnos serán resoltas de maneira individual no horario de titorías dos profesores da materia. Tamén se contemplan titorías en grupo.
-----------	--

## Avaliación

	Descrición	Cualificación	Resultados de Formación e Aprendizaxe
Resolución de problemas e/ou exercicios	Realización dun exame final teórico e/ou práctico	60	C1 C6 C8
Práctica de laboratorio	Realización de cuestionarios e actividades que serán expostas durante o período de docencia da materia.	40	C1 C2 C8 C10

## Outros comentarios sobre a Avaliación

Para superar a materia será necesario obter unha calificación mínima de 4 puntos (sobre 10) no exame final.

## Bibliografía. Fontes de información

## **Bibliografía Básica**

### **Bibliografía Complementaria**

Crawley, M. J., **Statistics: an introduction using R**, John Wiley and Sons, 2005

Crawley, M.J., **The R book**, John Wiley and Sons, 2013

Devore, Jay L., **Probability and statistics for engineering and sciences**, Cengage Learning, 2015

James, G., Witten, D., Hastie, T. and Tibshirani, R., **An Introduction to Statistical Learning with Applications in R**, Springer, 2013

Kabacoff, R., **R in Action: Data Analysis and Graphics with R**, Manning Publications, 2011

Maindonald, J. H., **Data analysis and graphics using R: an example-based approach**, Cambridge University Press., 2007

Tukey, J.W., **Exploratory Data Analysis**, Addison-Wesley, 1977

Zumel, N., Mount, J., **Practical Data Science with R**, Manning Publications, 2014

Zuur, A., Ieno, A., Meesters, E., **A Beginners` s Guide to R**, Springer, 2009

---

## **Recomendacións**

### **Outros comentarios**

Non se necesita cursar ningunha outra materia do máster. Con todo é fundamental a asistencia regular ás clases para a superación desta materia, xa que é moi importante o seguimento do traballo realizado na aula.

Os requisitos básicos desta materia son un coñecemento básico da Estatística e coñecementos a nivel usuario de \*Windows. Como xa se comentou utilizarase o software libre R.

Nalgunhas sesións realizaranse cuestionarios e exercicios sobre temas impartidos para a valoración da evolución e comprensión dos alumnos sobre a materia.

---

## **Plan de Continxencias**

### **Descrición**

No caso de que sexa necesario a importación de docencia en modalidade non presencial, a actividade docente impartirase mediante Campus Remoto ou outro plataforma semellante.

En calquera caso todo o material docente da materia será posto a disposición dos alumnos empregando algunha plataforma de intercambio de información (DropBox, plataforma de teledocencia Faitic, etc.)

De igual xeito, o exame final será feito de maneira non presencial, e non será necesario facer ningún cambio na planificación docente desta materia.

Ademais as titorías poderán realizarse por medios telemáticos (correo electrónico, videoconferencia, ...) baixo a modalidade de concertación previa.