

MÁSTER EN ECONOMÍA, FINANZAS Y COMPUTACIÓN • 2018-2019

1. DATOS BÁSICOS DE LA ASIGNATURA

ASIGNATURA (ESPAÑOL)	MINERÍA DE DATOS I		
SUBJECT	DATA MINING I		
CÓDIGO	1150205	AÑO DE PLAN DE ESTUDIOS	2015
TIPO	OBLIGATORIO	X	OPTATIVO
MÓDULO	Modelos predictivos y minería de datos I		
SEMESTRE	1º		
CRÉDITOS (ECTS)	3,0	TEORÍA (80%)	18 PRÁCTICAS (20%) 4,5

HORARIO DE CLASES

GRUPO	FECHA	Lunes	Martes	Miércoles	Jueves	Viernes
SESIÓN 1	09/01/2019			X		
SESIÓN 2	11/01/2019					X
SESIÓN 3	16/01/2019			X		
SESIÓN 4	18/01/2019					X
SESIÓN 5	23/01/2019			X		
SESIÓN 6	25/01/2019					X
SESIÓN 7	30/01/2019			X		
SESIÓN 8	06/02/2019			X		
SESIÓN 9	07/02/2019				X	
SESIÓN 10	08/02/2019					X

2. DOCENTES

RESPONSABLE DE LA ASIGNATURA

NOMBRE	DR. JOSÉ MARÍA MILLÁN (COORDINADOR)		
UNIVERSIDAD	HUELVA		
DEPARTAMENTO	ECONOMÍA		
ÁREA DE CONOCIMIENTO	ECONOMÍA APLICADA		
Nº DESPACHO	3	UBICACIÓN	FACULTAD DE CIENCIAS EMPRESARIALES Y TURISMO
CORREO ELECTRÓNICO	jose.millan@dege.uhu.es		TELÉFONO +34.959.217.886
URL WEB		CAMPUS VIRTUAL	Moodle

HORARIO DE TUTORÍAS (*)

La celebración de las tutorías se comunicará a los alumnos a través de la plataforma virtual de la asignatura. Se puede hacer uso de las tutorías asincrónicas a través de la plataforma y el correo electrónico. Para concertar citas presenciales, utilice el buzón de la asignatura con indicación de fecha y hora preferida.

OTROS DOCENTES

NOMBRE	DR. JOSÉ MARÍA ARRANZ MUÑOZ		
DEPARTAMENTO	ECONOMÍA		
UNIVERSIDAD	ALCALÁ DE HENARES		
ÁREA DE CONOCIMIENTO	ECONOMÍA APLICADA		
Nº DESPACHO	3	UBICACIÓN	SANTA MARÍA DE LA RÁBIDA
CORREO ELECTRÓNICO	josem.arranz@uah.es		TELÉFONO 918855198
URL WEB	https://portal.uah.es/portal/page/portal/epd2_profesores/prof121471	CAMPUS VIRTUAL	Moodle

HORARIO DE TUTORÍAS (*)

La celebración de las tutorías se comunicará a los alumnos a través de la plataforma virtual de la asignatura. Se puede hacer uso de las tutorías asincrónicas a través de la plataforma y el correo electrónico. Para concertar citas presenciales, utilice el buzón de la asignatura con indicación de fecha y hora preferida.

3. DESCRIPTOR

ESPAÑOL	Introducción al uso de STATA MP y Big Data con STATA. Estudio de las diferentes métodos y técnicas para la construcción de modelos predictivos y descriptivos a partir del análisis de bases de datos de alta cardinalidad y alta dimensionalidad, como ayuda a la toma de decisiones. Modelos de elección discreta. Modelos de supervivencia.
ENGLISH	Introduction to STATA MP and Big Data. The core of this subject on Data Mining is the study of different methods for analysing high dimensional data, and predictive models with discrete dependent variable. Survival models are also analysed.

4. SITUACIÓN

PRERREQUISITOS

Ninguno.

CONTEXTO DENTRO DE LA TITULACIÓN

Existen distintas asignaturas en el Master relacionadas con Modelos predictivos y data mining I. Especialmente complementaria con ella, son la de data mining II y Aprendizaje Automático, la primera continúa en la profundización de la materia con una orientación más especializada y la segunda aborda las técnicas de aprendizaje supervisado.

RECOMENDACIONES

La tónica general del curso será priorizar los conceptos clave y su aplicación práctica más que el puro formalismo teórico-matemático, de manera que solamente se necesitará un conocimiento básico de álgebra, estadística y cálculo.

5. COMPETENCIAS

BÁSICAS Y GENERALES

- CB9 - Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades
- CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

TRANSVERSALES

- CT1 – Dominar en un nivel intermedio una lengua extranjera
- CT2 - Utilizar de manera avanzada las tecnologías de la información y la comunicación.
- CT3 - Gestionar la información y el conocimiento.

ESPECÍFICAS

- CE2 - Comprender y saber aplicar los métodos de investigación cualitativa comúnmente utilizados en el ámbito de la Economía, la Empresa, las Finanzas y en el de la Comercialización e Investigación de mercados.
- CE4 – Conocer y saber utilizar el software comúnmente utilizados en el ámbito de la Economía, la Empresa, las Finanzas y el marketing
- CE5. Conocer los principios de programación informática y de gestión de bases de datos.

COMPETENCIA ESPECÍFICA ASOCIADA A LA ESPECIALIDAD

- CEF2 – Conocer y saber aplicar las técnicas más recientes en Finanzas Cuantitativas con especial

6. OBJETIVOS DE LA ASIGNATURA

- Conocimiento de los fundamentos teóricos y prácticos para el análisis de datos dentro del marco del aprendizaje no supervisado: Análisis de Clusters o Agrupaciones, Análisis de Componentes Principales, Análisis Factorial y Análisis, Análisis de componentes Independientes.
- Conocimiento de los fundamentos teóricos y prácticos para el análisis de datos dentro del marco del aprendizaje supervisado: Análisis Discriminante.
- Conocimiento de los fundamentos teóricos e instrumentales básicos para el análisis de modelos de variables dependientes discretas
- Conocimiento de los fundamentos teóricos e instrumentales básicos para el análisis de modelos de datos de duración.
- Adquisición y manejo de los términos específicos de la materia, teorías y aplicaciones básicas, conceptos elementales y adquisición de una visión global de su contenido con el programa Stata.
- Un acercamiento al campo para el estudio de tareas y técnicas, necesarias para la toma de decisiones de problemas complejos: problemas con solución inicial no algorítmica con datos estructurados, semiestructurados y no estructurados.
- El alumno ha de saber identificar y aplicar las técnicas más adecuadas a cada problema y plantearlos de forma adecuada.

7. ACTIVIDADES FORMATIVAS, METODOLOGÍA DE ENSEÑANZA-APRENDIZAJE Y SU RELACIÓN CON COMPETENCIAS QUE DEBE ADQUIRIR EL ESTUDIANTE

Metodología	Actividad	Descripción	Horas
Actividades presenciales	Clases teóricas	Clases teóricas. Fundamentos y planteamientos teóricos	25,5
	Clases prácticas	Problemas y casos prácticos: planteamiento y resolución de problemas concretos relacionados con la materia	
	Evaluación	Realización de exámenes parciales y finales, escritos u orales	
Trabajo autónomo tutelado	Trabajo autónomo individual	Uso de manuales, monografías y artículos (científicos, didácticos y divulgativos)	46,5
	Trabajo autónomo en grupo	Resolución de problemas y casos prácticos	
Tutorías	Individuales	Actividades académicas dirigidas	3
		Presenciales o virtuales (Campus Virtual, correo electrónico)	

Todo el material oportuno para el seguimiento de las clases teóricas y prácticas estará disponible en la plataforma de teleformación Moodle (<http://moodle.uhu.es/contenidos/login/index.php>) Para las clases teóricas, los recursos que se utilizarán son la pizarra (tradicional y en su versión electrónica), las proyecciones de presentaciones con la ayuda del ordenador y material suplementario suministrado por el profesorado (fotocopias, archivos electrónicos, etc.). En las clases prácticas se aplicarán los contenidos abordados en las clases teóricas, se hará hincapié en los mecanismos de resolución, sus limitaciones y ventajas, así como un análisis crítico de los resultados alcanzados. Estas clases prácticas serán interactivas y la participación del/la alumno/a será tenida en cuenta a la hora de valorar su adaptación al grado de aprendizaje.

8. BLOQUES TEMÁTICOS Y CONTENIDOS

Módulo I: Iniciación a STATA

Módulo II: Aprendizaje no supervisado: componentes principales; análisis factorial; análisis discriminante

Módulo III: Modelos de elección discreta

Módulo IV: Modelos de duración

9. BIBLIOGRAFÍA

BLIOGRAFÍA GENERAL

- Greene, W. (2011), *Econometric Analysis*, Ed. Prentice-Hall, 7ª Edición.
- James, G., Witten, D., Hastie, T., Tibshirani, R. (2013), *An Introduction to Statistical Learning with Applications in R*, Springer
- Koch, I. (2013), *Analysis of Multivariate and High-Dimensional Data*, Cambridge.
- Rencher, A.C., Christensen, W.F. (2012), *Methods of Multivariate Analysis*, 3rd Edition, Wiley.
- Wooldridge, J. (2010), *Econometric Analysis of Cross Section and Panel Data*. Ed. MIT Press, 2ª Edición.

BIBLIOGRAFÍA COMPLEMENTARIA

- Adkins L. C. y R. C. Hill (2011), *Using Stata for Principles of Econometrics*. Ed. John Wiley & Sons
- Cameron C, y P. Trivedi, (2009), *Microeconometrics using Stata*. Ed. Stata Press.
- Hair, J.F., Black, W.C, Babin, B.J., and Anderson, R.E. (2010), *Multivariate Data Analysis*, 7th Edition, Prentice Hall, Inc
- Wooldridge, J., (2010), *Introducción a la Econometría. Un enfoque moderno*. Ed. Cengage Learning, 4a Edición.
- Wooldridge, J., (2003), *Solutions manual and supplementary materials for Econometric Analysis of Cross Section and Panel Data*. Ed. MIT Press.

10. SISTEMA DE EVALUACIÓN

Técnica empleada	Descripción	Criterios	Valor sobre el total de la nota
Prueba final individual teórico-práctica	Evaluación de la adquisición conocimientos teórico-prácticos y metodológicos	<ul style="list-style-type: none"> • Grado de capacidad de resolución de problemas y aplicación de los contenidos teóricos a la práctica • Grado de desarrollo de la capacidad de síntesis • Grado de conocimiento, comprensión e información • Ausencia de errores • Utilización adecuada de los conceptos • Coherencia interna del ejercicio • Capacidad de interrelacionar teorías, modelos, conceptos • Concreción y exactitud de las respuestas • Nivel de estudio 	50%
Evaluación continua	Valoración del trabajo personal a través de portafolios (ejercicios prácticos realizados tanto de manera autónoma como en grupo) Participación activa	<ul style="list-style-type: none"> • Capacidad del alumnado para planificar, desarrollar y presentar un trabajo empírico sobre diferentes facetas de la asignatura • Claridad de análisis y exposición de resultados • Grado de capacidad en la resolución de problemas 	50%

Por defecto, la evaluación de la asignatura resultará de sumar la nota de la prueba objetiva final, a celebrar cuando marque el calendario de exámenes de la Facultad (que puede incluir la elaboración de un programa, cuestiones de elección múltiple y/o cuestiones teórico-prácticas), con un valor máximo de cinco puntos, y la puntuación obtenida en el portafolio entregado por el alumno (máximo 5 puntos) a través de la plataforma virtual en las fechas señaladas, a través de tests de resultados y otras pruebas prácticas. Las calificaciones correspondientes a estas actividades se conservarán de cara a la nota final en las convocatorias ordinarias I y II. Sin perjuicio de lo anterior, los alumnos que así lo deseen pueden solicitar en tiempo y forma según está recogido en el Reglamento Evaluación para las Titulaciones de Grado y Máster Oficial de la Universidad de Huelva, la evaluación única final. Esta prueba única final consta de un examen con cuestiones de elección múltiple y/o cuestiones teórico-prácticas acerca del total del programa incluido en la presente guía docente, que supone el 100% de la calificación de la asignatura. En la convocatoria ordinaria III la evaluación de la asignatura se realizará para todo el alumnado en base a un único examen con cuestiones de elección múltiple y/o cuestiones teórico-prácticas acerca del total del programa incluido en la presente guía docente, que supone el 100% de la calificación de la asignatura.

Los criterios de evaluación y calificación serán los que marca Reglamento Evaluación para las Titulaciones de Grado y Máster Oficial de la Universidad de Huelva, valorándose la capacidad de comprensión y relación, la capacidad de síntesis, la actitud crítica, la capacidad y profundidad de análisis y aplicación de los modelos, la originalidad, la relación entre conceptos teóricos y aplicaciones y la utilización de las fórmulas y modelos adecuados en los ejercicios numéricos. El conjunto de las actividades de evaluación estará sujeto al Reglamento de Evaluación para las Titulaciones de Grado y Máster Oficial de la Universidad de Huelva.

El sistema de calificación empleado en la materia está de acuerdo con el establecido en artículo 5 del Real Decreto 1125/2003, de 5 de septiembre, por el que se establece el sistema europeo de créditos y el sistema de calificaciones en las titulaciones universitarias de carácter oficial y de validez en todo el territorio nacional: Los resultados obtenidos por el/la alumno/a en cada una de las materias del plan de estudios se calificarán en función de la siguiente escala numérica de 0 a 10, con expresión de un decimal, a la que podrá añadirse su correspondiente calificación cualitativa:

- 0,0 a 4,9: Suspenso (SS)
- 5,0 a 6,9: Aprobado (AP)

- 7,0 a 8,9: Notable (NT)
- 9,0 a 10: Sobresaliente (SB)

La mención "Matrícula de Honor" podrá ser otorgada a alumnos/as que hayan obtenido una calificación igual o superior a 9,0. Su número no podrá exceder del 5% de los/las alumnos/as matriculados/as en una materia en el correspondiente curso académico, salvo que el número de alumnos/as matriculados/as sea inferior a 20, en cuyo caso se podrá conceder una sola "Matrícula de Honor". En caso de que haya más candidatos que posibilidades de matrículas de honor por número de estudiantes en la asignatura, se otorgará la matrícula de honor a aquel alumno con mayor calificación en el examen final.

MEDIDAS PREVISTAS PARA RESPONDER A NECESIDADES EDUCATIVAS ESPECIALES

Se adoptarán, las medidas adecuadas a cada caso para que aquellos alumnos que presenten necesidades especiales puedan adquirir los conocimientos y capacidades necesarias para la superación de la materia.

D) ORGANIZACIÓN DEL TRABAJO DURANTE EL CURSO

11. NÚMERO DE HORAS DE TRABAJO DEL ESTUDIANTE

Nº de Horas: 75 (3 créditos ECTS)

- Actividades presenciales: 25,5 horas
 - Clases de aula teóricas: Método expositivo. 18 horas
 - Clases de aula de problemas: Método expositivo. 4,5 horas
 - Sesiones de evaluación: 3 horas
- Trabajo autónomo tutelado: 46,5 horas
 - Trabajo autónomo individual: 36,5 horas
 - Trabajo autónomo en grupo: 10 horas
 - Tutorías docentes: 3 horas

E) TEMARIO DESARROLLADO

Tema.1. Iniciación al uso de STATA

- 1.1. Ventanas y ficheros
- 1.2. Preparando STATA para trabajar
- 1.3. Cargando bases de datos
- 1.4. Estructura básica de los comandos
- 1.5. Análisis descriptivo de datos
- 1.6. Gestión de variables
- 1.7. Fusiones verticales y horizontales: los comandos 'append' y 'merge'
- 1.8. Generación de retardos y diferencias
- 1.9. Estimación: MCO
- 1.10. La utilización de STATA para el Big Data

Tema 2. Aprendizaje no supervisado

- 2.1. Análisis de conglomerados
- 2.2. Análisis de componentes principales
- 2.3. Análisis factorial

Tema 3. Modelos de elección binaria

- 3.1. Modelos binarios
- 3.2. Modelos de probabilidad lineal
- 3.3. Modelos probit y logit
- 3.4. Ejemplos de aplicación.

Tema 4. Modelos de Supervivencia

- 4.1. Introducción
- 4.2. Herramientas de los modelos de duración
- 4.3. Modelos de duración de tiempo continuo y discreto
- 4.4. Métodos de estimación.

F) MECANISMOS DE CONTROL Y SEGUIMIENTO DE LA ASIGNATURA

Cada vez que finalice un módulo del programa se realizarán pruebas diversas para establecer el nivel de captación de competencias y contenidos de la asignatura.