

## 1. DATOS BÁSICOS

<b>Asignatura</b>	Ingeniería de datos I
<b>Titulación</b>	Grado en gestión empresarial basada en el análisis de datos
<b>Escuela/ Facultad</b>	CC. Sociales y Comunicación
<b>Curso</b>	Segundo
<b>ECTS</b>	6
<b>Carácter</b>	Obligatoria
<b>Idioma/s</b>	Castellano
<b>Modalidad</b>	Presencial
<b>Semestre</b>	Segundo
<b>Curso académico</b>	2024-25
<b>Docente coordinador</b>	Hugo Fernandez Visier

## 2. PRESENTACIÓN

Ingeniería de Datos, enfocada en el tratamiento de grandes volúmenes de información. Así, se analizarán los modos de agregar la información, como las bases de datos distribuidas, los data warehouses o las bases de datos en la nube (cloud databases). Al último nivel se encuentra el enfoque en el Big Data, cuyo objetivo es la gestión de enormes volúmenes de información, lo cual requiere arquitecturas desestructuradas en las que prime la eficiencia en el acceso

## 3. COMPETENCIAS Y RESULTADOS DE APRENDIZAJE

### Competencias básicas:

- CB4 - Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.

### Competencias transversales:

- CT2 - Aprendizaje autónomo: Conjunto de habilidades para seleccionar estrategias de búsqueda, análisis, evaluación y gestión de la información procedente de fuentes diversas, así como para aprender y poner en práctica de manera independiente lo aprendido.
- CT4 - Comunicación escrita / Comunicación oral: Capacidad para transmitir y recibir datos, ideas, opiniones y actitudes para lograr comprensión y acción, siendo oral la que se realiza mediante palabras y gestos y, escrita, mediante la escritura y/o los apoyos gráficos.

### Competencias específicas:

- CE20 - Capacidad para desarrollar un plan de estudio de variables e indicadores empresariales.

- CE21 - Capacidad para segmentar y seleccionar el público objetivo adecuado a los intereses de cada iniciativa empresarial
- CE22 - Capacidad para seleccionar y aplicar las herramientas analíticas más adecuadas a cada situación de la compañía
- CE23 - Capacidad para seleccionar, parametrizar y analizar métricas relacionadas con las fuentes de información en función de cada ámbito (marketing, finanzas, etc.)
- CE24 - Capacidad para identificar las necesidades de información, dependiendo de los objetivos marcados
- CE26 - Espíritu crítico y objetividad para cuestionar los datos o las verdades asumidas a partir de los datos precedentes
- CE28 - Capacidad para entender y conocer el "ciclo del dato": adquisición y creación de datos, construcción de la información, análisis y visualización
- CE29 - Capacidad para hacer las preguntas correctas, en relación con el objetivo de conocimiento esperado, para que se traduzcan en las "queries" adecuadas que se deben formular al sistema de almacenamiento de datos
- CE31 - Capacidad para gestionar la incertidumbre debida al constante cambio de las fuentes de información
- CE33 - Capacidad para manejar con soltura y solvencia técnica herramientas informáticas de tratamiento estadístico y otras como simuladores

#### Resultados de aprendizaje:

- RA1. Conocer algunas de las principales bases de datos utilizadas y tecnologías utilizadas para el almacenamiento de datos
- RA2. Conocer la importancia de la ciencia de datos
- RA3. Describir las técnicas de aprendizaje automático, seleccionar la más adecuada y diseñar una solución a un problema dado que las utilice.
- RA4. Aprender como tratar el problema de la alta dimensionalidad en el ámbito del Big Data conectándolas con casos reales y argumentar su evolución y aplicaciones futuras.
- RA5. Conocer cuando se utilizan técnicas de aprendizaje supervisado o no supervisado.
- RA6. Utilización de la la visualización de los datos para mejorar la interpretabilidad de los análisis haciendo uso de los lenguajes y entornos específicos.
- RA7. Conocer diferentes casos de uso de las técnicas de aprendizaje automático en el ámbito empresarial y de negocio
- RA8. Conocer cuales son los principales problemas que se puede encontrar un analista de datos y como enfrentarlos

En la tabla inferior se muestra la relación entre las competencias que se desarrollan en la asignatura y los resultados de aprendizaje que se persiguen:

Competencias	Resultados de aprendizaje
CB4, CT2, CT4, CE31	RA1, RA2
CE24, CE29, CE31	RA3
CE22, CE25	RA4, RA5
CE20, CE21, CE23, CE28, CE26, CE29, CE33, CE31	RA6, RA7, RA8

## 4. CONTENIDOS

- Fundamentos de la gestión de grandes volúmenes de datos
- Bases de datos no convencionales. NoSQL.
- Tipos de bases de datos noSQL y datos semiestructurados
- Introducción a Python y librerías de ingeniería de datos
- Introducción a la clasificación supervisada y no supervisada
- Técnicas exploratorias de análisis (clustering)
- Procesos de minería de datos
- Metodologías de validación en el aprendizaje
- Introducción a sistemas lineales de aprendizaje: casos de estudio sobre regresión lineal y regresión logística
- Selección y extracción de características: casos de estudio sobre técnicas de reducción de la dimensionalidad (PCA, LDA)
- Aprendizaje inductivo y árboles de decisión: caso de estudio sobre generación de reglas automáticamente
- Introducción al aprendizaje profundo

## 5. METODOLOGÍAS DE ENSEÑANZA-APRENDIZAJE

A continuación, se indican los tipos de metodologías de enseñanza-aprendizaje que se aplicarán:

- Clase magistral/ web conference
- Método del caso
- Aprendizaje basado en problemas (ABP)
- Aprendizaje basado en enseñanzas de taller
- 

## 6. ACTIVIDADES FORMATIVAS

A continuación, se identifican los tipos de actividades formativas que se realizarán y la dedicación en horas del estudiante a cada una de ellas:

Actividad formativa	Número de horas
Tutorías	10
Lecciones magistrales	20
Lecciones magistrales asíncronas	10
Análisis de casos	15
Resolución de problemas	20
Elaboración de informes y escritos	15
Trabajo autónomo	30
Actividades en talleres y/o laboratorios	30

<b>TOTAL</b>	<b>150</b>
--------------	------------

## 7. EVALUACIÓN

A continuación, se relacionan los sistemas de evaluación, así como su peso sobre la calificación total de la asignatura:

Sistema de evaluación	Peso
Pruebas presenciales de conocimiento	30.0
Exposiciones orales	20.0
Caso/problema	25.0
Práctica de laboratorio	25.0

En el Campus Virtual, cuando accedas a la asignatura, podrás consultar en detalle las actividades de evaluación que debes realizar, así como las fechas de entrega y los procedimientos de evaluación de cada una de ellas.

### 7.1. Convocatoria ordinaria

Para superar la asignatura en convocatoria ordinaria deberás obtener una calificación mayor o igual que 5,0 sobre 10,0 en la calificación final (media ponderada) de la asignatura.

En todo caso, será necesario que obtengas una calificación mayor o igual que 4,0 en la prueba final, para que la misma pueda hacer media con el resto de las actividades.

### 7.2. Convocatoria extraordinaria

Para superar la asignatura en convocatoria ordinaria deberás obtener una calificación mayor o igual que 5,0 sobre 10,0 en la calificación final (media ponderada) de la asignatura.

En todo caso, será necesario que obtengas una calificación mayor o igual que 4,0 en la prueba final, para que la misma pueda hacer media con el resto de las actividades.

Se deben entregar las actividades no superadas en convocatoria ordinaria, tras haber recibido las correcciones correspondientes a las mismas por parte del docente, o bien aquellas que no fueron entregadas.

## 8. CRONOGRAMA

En este apartado se indica el cronograma con fechas de entrega de actividades evaluables de la asignatura:

Actividades evaluables	Fecha
Actividad 1. Introducción a librerías de manejo de lectura de datos e interpretación de datos	Semana 2
Actividad 2. Acceso a bases de datos No-SQL y manejo de datos	Semana 5
Actividad 3. Selección de variables importantes	Semana 8
Actividad 4. Modelos no supervisados	Semana 11
Actividad 5. Modelos supervisados	Semana 13
Actividad 6. Practica Pyspark	Semana 15
Prueba final.	Semana 18

Este cronograma podrá sufrir modificaciones por razones logísticas de las actividades. Cualquier modificación será notificada al estudiante en tiempo y forma.

## 9. BIBLIOGRAFÍA

- **"Python for Data Analysis" by Wes McKinney:** Este libro escrito por el creador de Pandas, Wes McKinney, es una excelente introducción al preprocesamiento de datos con Pandas. Libro: [Python for Data Analysis](#)
- **Pandas Documentation:** La documentación oficial de Pandas es un recurso esencial para aprender sobre la biblioteca y cómo preprocesar datos utilizando sus funciones y métodos. Sitio web: [Pandas Documentation](#)
- **MongoDB University:** MongoDB University ofrece cursos en línea gratuitos sobre MongoDB, una popular base de datos NoSQL Documental. Sitio web: [MongoDB University](#)
- **Cassandra Documentation:** La documentación oficial de Apache Cassandra, una base de datos NoSQL altamente escalable y distribuida. Sitio web: [Cassandra Documentation](#)
- **Neo4j Learning Center:** Neo4j es una base de datos de gráficos y su Learning Center ofrece recursos para aprender sobre bases de datos de gráficos y Cypher (lenguaje de consulta). Sitio web: [Neo4j Learning Center](#)
- **"Introduction to Machine Learning with Python" by Andreas C. Müller & Sarah Guido:** Este libro se centra en el uso de Python y scikit-learn para aprender machine learning. Es una excelente introducción práctica. Libro: [Introduction to Machine Learning with Python](#)
- **"Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" by Aurélien Géron:** Un libro práctico que cubre el uso de scikit-learn, Keras y TensorFlow para aplicaciones de machine learning. Libro: [Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow](#)
- **Documentación oficial de scikit-learn:** La documentación oficial es el recurso más completo para aprender sobre scikit-learn, con tutoriales, ejemplos y una descripción detallada de todas las funcionalidades. Sitio web: [Scikit-learn Documentation](#)

## 10. UNIDAD DE ATENCIÓN A LA DIVERSIDAD

Desde la Unidad de Orientación Educativa y Diversidad (ODI) ofrecemos acompañamiento a nuestros estudiantes a lo largo de su vida universitaria para ayudarles a alcanzar sus logros académicos. Otros de los pilares de nuestra actuación son la inclusión del estudiante con necesidades específicas de apoyo educativo, la accesibilidad universal en los distintos campus de la universidad y la equiparación de oportunidades.

Desde esta Unidad se ofrece a los estudiantes:

1. Acompañamiento y seguimiento mediante la realización de asesorías y planes personalizados a estudiantes que necesitan mejorar su rendimiento académico.
2. En materia de atención a la diversidad, se realizan ajustes curriculares no significativos, es decir, a nivel de metodología y evaluación, en aquellos alumnos con necesidades específicas de apoyo educativo persiguiendo con ello una equidad de oportunidades para todos los estudiantes.
3. Ofrecemos a los estudiantes diferentes recursos formativos extracurriculares para desarrollar diversas competencias que les enriquecerán en su desarrollo personal y profesional.
4. Orientación vocacional mediante la dotación de herramientas y asesorías a estudiantes con dudas vocacionales o que creen que se han equivocado en la elección de la titulación.

Los estudiantes que necesiten apoyo educativo pueden escribirnos a:

[orientacioneducativa@universidadeuropea.es](mailto:orientacioneducativa@universidadeuropea.es)

## 11. ENCUESTAS DE SATISFACCIÓN

¡Tú opinión importa!

La Universidad Europea te anima a participar en las encuestas de satisfacción para detectar puntos fuertes y áreas de mejora sobre el profesorado, la titulación y el proceso de enseñanza-aprendizaje.

Las encuestas estarán disponibles en el espacio de encuestas de tu campus virtual o a través de tu correo electrónico.

Tu valoración es necesaria para mejorar la calidad de la titulación.

Muchas gracias por tu participación.