

## Ingeniería de Datos

Código: 106565  
Créditos ECTS: 6

**2025/2026**

Titulación	Tipo	Curso
Inteligencia Artificial / Artificial Intelligence	OB	1

### Contacto

Nombre: Javier Vazquez Corral

Correo electrónico: [javier.vazquez.corral@uab.cat](mailto:javier.vazquez.corral@uab.cat)

### Idiomas de los grupos

Puede consultar esta información al [final](#) del documento.

### Prerrequisitos

Los estudiantes deberán haber cursado y entendido las asignaturas: "Introduction to Programming I" i "Mathematical Foundations I".

### Objetivos y contextualización

La asignatura pretende proporcionar los fundamentos del análisis y visualización de datos. Se estudiarán las diferentes etapas de los procesos de análisis de datos, desde la recolección, anotación y preparación de datos, hasta su análisis y visualización, preparando el camino para un modelado más avanzado a través de Machine Learning (M8).

### Competencias

- Conceptualizar y modelar alternativas de soluciones complejas a problemas de aplicación de la inteligencia artificial en diferentes ámbitos, y planificar y gestionar proyectos para el diseño y desarrollo de prototipos que demuestren la validez del sistema propuesto.
- Conocer, comprender, utilizar y aplicar de forma adecuada los fundamentos matemáticos necesarios para desarrollar sistemas de razonamiento, aprendizaje y manipulación de grandes volúmenes de datos.
- Conocer y utilizar de forma eficiente las técnicas y herramientas de representación, manipulación, análisis y gestión de datos a gran escala.
- Introducir cambios en los métodos y los procesos del ámbito de conocimiento para dar respuestas innovadoras a las necesidades y demandas de la sociedad.
- Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.
- Trabajar cooperativamente para la consecución de objetivos comunes, asumiendo la propia responsabilidad y respetando el rol de los diferentes miembros del equipo.

### Resultados de aprendizaje

1. Concebir, diseñar e implementar procesos de recopilación y anotación de datos adecuados al problema concreto a resolver.
2. Conocer las herramientas básicas de manipulación de diferentes tipos de datos estructurados, semiestructurados y no estructurados.
3. Proponer nuevos métodos o soluciones alternativas fundamentadas.
4. Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado.
5. Trabajar cooperativamente para la consecución de objetivos comunes, asumiendo la propia responsabilidad y respetando el rol de los diferentes miembros del equipo.
6. Utilizar adecuadamente los métodos de visualización de datos.
7. Utilizar métodos de análisis de datos para probar hipótesis, y obtener interpretaciones útiles.

## Contenido

La asignatura se divide en 11 temas principales:

- Introducción.
- Repaso de estadística básica.
- Numpy / Matplotlib/ Pandas
- XML/JSON
- Expresiones regulares/DFA/NFA
- Tipos de datos. Datos que faltan.
- Introducción al análisis de datos. Outliers.
- PCA
- kNN
- Image retrieval y recommender systems.
- K-means

## Actividades formativas y Metodología

Título	Horas	ECTS	Resultados de aprendizaje
<b>Tipo: Dirigidas</b>			
Clases de ejercicios	12	0,48	2, 6, 7
Clases de teoría	28	1,12	2, 3, 7
<b>Tipo: Supervisadas</b>			
Clases de laboratorio	12	0,48	1, 2, 3, 4, 5, 6, 7
<b>Tipo: Autónomas</b>			
Completación de los ejercicios en casa	16	0,64	

Estudio de la teoría	20	0,8	2
Realización de los proyectos de prácticas	58	2,32	1, 2, 3, 4, 5, 6, 7

Habrá tres tipos de sesiones:

Clases de teoría: El objetivo de estas sesiones es que el profesor explique los antecedentes teóricos de la materia. Para cada uno de los temas estudiados se explica la teoría y formulación matemática, así como las soluciones algorítmicas correspondientes.

Sesiones de ejercicios: Serán sesiones que faciliten la interacción. En estas sesiones se pretende reforzar la comprensión de los temas vistos en las clases de teoría proponiendo casos prácticos que requieran el diseño de una solución en la que se utilicen los métodos vistos en las clases de teoría.

Sesiones prácticas de laboratorio: Serán sesiones en las que se realizarán diferentes tipos de actividades relacionadas con la realización del proyecto/proyectos por grupos de alumnos. Durante las sesiones prácticas se presentará el proyecto/proyectos a resolver y se realizarán una serie de actividades en grupo en modalidad de trabajo colaborativo. Se abordará la identificación del problema, la discusión del diseño, la distribución y organización del trabajo a realizar, el desarrollo de la solución y la presentación de los resultados al profesor y al resto de estudiantes.

Toda la información de la asignatura y los documentos relacionados que necesiten los estudiantes se encontrarán en el campus virtual.

Dentro del horario marcado por el centro o titulación, se reservarán 15 minutos de una clase para que los estudiantes evalúen a su profesorado y sus cursos o módulos a través de cuestionarios.

Uso de la IA: Para esta asignatura, se permite el uso de tecnologías de Inteligencia Artificial (IA) exclusivamente en tareas de apoyo, como la búsqueda bibliográfica o de información, la corrección de textos o las traducciones. El estudiante deberá identificar claramente qué partes han sido generadas con esta tecnología, especificar las herramientas utilizadas e incluir una reflexión crítica sobre cómo estas han influido en el proceso y el resultado final de la actividad. La falta de transparencia en el uso de la IA en esta actividad evaluable se considerará una falta de honestidad académica y puede conllevar una penalización parcial o total en la nota de la actividad, o sanciones mayores en casos graves.

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

## Evaluación

### Actividades de evaluación continuada

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Examen Parcial 1	0.2	2	0,08	1, 2, 3, 4, 7
Examen Parcial 2	0.2	2	0,08	1, 2, 3, 4, 7
Portfolio	0.2	0	0	1, 2, 3, 4, 6, 7
Proyecto de prácticas	0.4	0	0	1, 2, 3, 4, 5, 6, 7

Esta asignatura no prevee el sistema de evaluación única.

Nota final:

La nota final se calcula mediante una media ponderada en función de las diferentes actividades que se realizan:

$$\text{Nota final} = 0,4 * \text{Nota de teoría} + 0,2 * \text{Nota de ejercicios} + 0,4 * \text{Nota de laboratorio}$$

Para aplicar esta fórmula, la condición es que tanto la nota de teoría como la de laboratorio sean superiores a 5. Cuando un alumno no alcance el mínimo exigido en alguna de las actividades de evaluación, la nota final será la del elemento no permitiendo el cálculo (es decir, si un estudiante tiene un 6 en Teoría, un 5 en Ejercicios, pero un 2 en Laboratorio, la nota final será un 2). La calificación de "no evaluable" se dará a las personas de las que no se tenga ninguna entrega.

Nota de teoría

La nota de teoría tiene como objetivo evaluar las capacidades individuales del estudiante en cuanto a los contenidos teóricos de la asignatura. Esto se hace de forma continua durante el curso a través de dos exámenes parciales:

$$\text{Nota de Teoría} = 0,5 * \text{Nota Examen 1} + 0,5 * \text{Nota Examen 2}$$

El examen 1 se realiza a mitad del cuatrimestre y sirve para eliminar parte de la materia si se aprueba. El examen 2 se realiza al final del cuatrimestre y sirve para eliminar parte de la materia si se aprueba. Estos exámenes tienen como objetivo evaluar las habilidades de cada alumno de forma individualizada, tanto en lo que se refiere a la resolución de ejercicios utilizando las técnicas explicadas en clase, como a evaluar el nivel de conceptualización que el alumno ha realizado de las técnicas vistas. Para obtener una nota final de teoría aprobada, se requerirá que las notas 1 y 2 del examen parcial sean ambas superiores a 4.

Examen de recuperación. En caso de que la nota de teoría no alcance el nivel adecuado para aprobar, los alumnos podrán realizar un examen de recuperación, destinado a recuperar la parte reprobada (1, 2 o ambas) del proceso de evaluación continua. Nota: El examen de recuperación también se puede realizar si el alumno desea obtener una calificación superior a las anteriores. Dicho esto, este examen no permitirá la obtención de la calificación "Matrícula de Honor".

Nota de ejercicios:

El objetivo de los ejercicios es que el estudiante se entrene con los contenidos de la materia de forma continuada y se familiarice con la aplicación de los conceptos teóricos. Como prueba de este trabajo se solicita la presentación de un portafolio en el que se cotejarán los ejercicios realizados:

Nota de ejercicios = Evaluación del portfolio

Nota de Prácticas de Laboratorio:

<span class="Q4iAWc"> La parte de prácticas de laboratorio tiene un peso fundamental en la nota global de la asignatura. Las sesiones de laboratorio tienen como objetivo que el estudiante diseñe una solución a un problema planteado de forma contextualizada. Tales problemas requerirán el diseño de una solución integral, desde la exploración de las técnicas disponibles hasta el modelado de datos. Además, los estudiantes deben demostrar sus habilidades de trabajo en equipo y presentar los resultados a la clase de manera convincente. Las sesiones de laboratorio se estructuran en torno a proyecto/s. Los proyectos se evalúan a través de su entregable, una exposición oral que los alumnos realizarán en clase y un proceso de autoevaluación. La calificación se calcula de la siguiente manera:

$$\text{Calificación del proyecto} = 0,6 * \text{Calificación de los entregables} + 0,3 * \text{Calificación de la presentación} + 0,1 * \text{Calificación de la autoevaluación}$$

En caso de ser más de 1 proyecto, la Nota del Laboratorio será la media de todas las Notas de los Diferentes Proyectos. En caso de no aprobar el proyecto, se permitirá la recuperación de la parte de los entregables de

los proyectos no aprobados, restringida a una nota máxima de 7/10. La presentación oral no se puede recuperar.

## Bibliografía

Data Science from Scratch: First Principles with Python, Joel Grus, O'Reilly Media, 2015, 1st Ed.

Python Data Science Handbook, Jake Van der Plas, O'Reilly Media, 2016, 1st Ed.

Computational and Inferential Thinking: The Foundations of Data Science, Ani Adhikari and John DeNero, online: <https://ds8.gitbooks.io/textbook/content/>

## Software

Python

## Grupos e idiomas de la asignatura

La información proporcionada es provisional hasta el 30 de noviembre de 2025. A partir de esta fecha, podrá consultar el idioma de cada grupo a través de este [enlace](#). Para acceder a la información, será necesario introducir el CÓDIGO de la asignatura

Nombre	Grupo	Idioma	Semestre	Turno
(PAUL) Prácticas de aula	71	Inglés	segundo cuatrimestre	tarde
(PLAB) Prácticas de laboratorio	1	Inglés	segundo cuatrimestre	tarde
(PLAB) Prácticas de laboratorio	2	Inglés	segundo cuatrimestre	tarde
(TE) Teoría	1	Inglés	segundo cuatrimestre	tarde