



UNIVERSIDAD  
POLITÉCNICA  
DE MADRID

PROCESO DE  
COORDINACIÓN DE LAS  
ENSEÑANZAS PR/CL/001



E.T.S. de Ingeniería de  
Sistemas Informáticos

# ANX-PR/CL/001-01

## GUÍA DE APRENDIZAJE

### ASIGNATURA

**615001030 - Bases De Datos Ii**

### PLAN DE ESTUDIOS

61CD - Grado En Ciencia De Datos E Inteligencia Artificial

### CURSO ACADÉMICO Y SEMESTRE

2022/23 - Segundo semestre

## Índice

---

### Guía de Aprendizaje

1. Datos descriptivos.....	1
2. Profesorado.....	1
3. Conocimientos previos recomendados.....	2
4. Competencias y resultados de aprendizaje.....	2
5. Descripción de la asignatura y temario.....	3
6. Cronograma.....	5
7. Actividades y criterios de evaluación.....	7
8. Recursos didácticos.....	9
9. Otra información.....	10

## 1. Datos descriptivos

---

### 1.1. Datos de la asignatura

<b>Nombre de la asignatura</b>	615001030 - Bases de Datos II
<b>No de créditos</b>	6 ECTS
<b>Carácter</b>	Obligatoria
<b>Curso</b>	Tercero curso
<b>Semestre</b>	Sexto semestre
<b>Período de impartición</b>	Febrero-Junio
<b>Idioma de impartición</b>	Castellano
<b>Titulación</b>	61CD - Grado en Ciencia de Datos e Inteligencia Artificial
<b>Centro responsable de la titulación</b>	61 - Escuela Tecnica Superior De Ingenieria De Sistemas Informaticos
<b>Curso académico</b>	2022-23

## 2. Profesorado

---

### 2.1. Profesorado implicado en la docencia

<b>Nombre</b>	<b>Despacho</b>	<b>Correo electrónico</b>	<b>Horario de tutorías</b> *
Edgar Talavera Muñoz (Coordinador/a)	1222	e.talavera@upm.es	Sin horario.
Alberto Diaz Alvarez	4122	alberto.diaz@upm.es	Sin horario.
Angel Panizo Lledot	1214	angel.panizo@upm.es	Sin horario.

\* Las horas de tutoría son orientativas y pueden sufrir modificaciones. Se deberá confirmar los horarios de tutorías con el profesorado.

## 3. Conocimientos previos recomendados

---

### 3.1. Asignaturas previas que se recomienda haber cursado

- Fundamentos De La Programación
- Bases De Datos I
- Programación Para Ciencia De Datos

### 3.2. Otros conocimientos previos recomendados para cursar la asignatura

- Bases de datos relacionales
- Programación

## 4. Competencias y resultados de aprendizaje

---

### 4.1. Competencias

CB03 - Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética

CB04 - Que los estudiantes puedan transmitir información, ideas, problemas y soluciones a un público tanto especializado como no especializado

CE05 - Capacidad de diseñar e implementar los procesos de selección, limpieza, transformación, integración y verificación de la calidad de los datos de cara a su posterior tratamiento.

CE07 - Capacidad de diseñar e implementar sistemas de información (incluyendo modelos de datos y estrategias de gestión de datos) dimensionados para gestionar el volumen, velocidad y variedad de los datos, de forma adecuada para su almacenamiento, procesamiento y acceso para tratamientos posteriores.

## 4.2. Resultados del aprendizaje

RA129 - Ser capaz de implementar y gestionar una base de datos en un gestor no relacional

RA89 - Usar lenguajes de programación y de descripción de datos, comunes en Ciencia de Datos

RA76 - RA-APID-5 Configuración, administración, uso y optimización de sistemas gestores de bases de datos relacionales

RA77 - RA-APID-6 Diseño, creación, consulta y manipulación de repositorios de datos, e integración con aplicaciones del sistema

RA134 - RA114 - RA-APID-18 Ser capaz de utilizar las tecnologías de información para preparar los conjuntos de datos

## 5. Descripción de la asignatura y temario

---

### 5.1. Descripción de la asignatura

El objetivo de la asignatura "Bases de Datos II" es dar continuidad al alumno en el conocimiento de diferentes tipos de sistemas de almacenamiento de información. Tras haber cursado y visto bases de datos, donde se focaliza la asignatura en las bases de datos relacionales, esta asignatura pretende cubrir en primer lugar, otro tipología de bases de datos que siguen el modelo relacional como son las bases de datos basadas en ficheros (Que no requieren de instalaciones de SGBDs para su manejo y que permiten construir bases de datos ligeras de forma sencilla) y las bases de datos federadas, donde se plantea la posibilidad de generación de una estructura de bases de datos dispersas para que funcionen como si fueran una sola entidad.

La asignatura continua en el bloque 2 con los conceptos de gobernanza de datos y provenance, mostrando al alumno fundamentos teóricos que deben tenerse en cuenta a la hora de manejos de datos, y sobre todo de aquellos metadatos asociados a las fuentes de información. Se completa este tema con componentes prácticos referidos a la curación y limpieza de datos, así como al desarrollo de procesos de extracción y transformación de información mediante uso de lenguajes de programación que permitan transformar tipologías de datos y serializarlos en formatos diferentes según el caso de uso necesario.

El bloque 3 entra en las bases de datos NoSQL, empezando con los fundamentos teóricos en los que se sustentan estas bases de datos, pero mostrando enseguida un enfoque práctico, revisando diferentes tipos de bases de datos NoSQL y tratando de mostrar al alumno como acceder a estas desde un punto de vista programático.

Finalmente el bloque 4 se centra en el concepto de data streaming y el manejo de datos que puedan ser generados en tiempo real y que necesitan de arquitecturas específicas.

## 5.2. Temario de la asignatura

1. Bloque I: Bases de datos avanzadas
  - 1.1. Bases de datos basadas en ficheros
  - 1.2. Bases de datos federadas
2. Bloque II: Gobernanza de datos
  - 2.1. Gobernanza de datos
  - 2.2. Data provenance
  - 2.3. Curación y limpieza de datos
  - 2.4. Transformación y serialización de datos
3. Bloque III: Bases de datos NoSQL
  - 3.1. Fundamentos de bases de datos NoSQL
  - 3.2. Comparativa de bases de datos relacionales vs NoSQL
  - 3.3. Tipos y características de bases de datos NoSQL
  - 3.4. Bases de datos documentales
  - 3.5. Bases de datos basadas en grafos
  - 3.6. Bases de datos basadas en clave-valor
  - 3.7. Otras bases de datos (dispersas, columnares, en memoria, ...)
4. Bloque IV: Data streaming
  - 4.1. Procesamiento de datos continuos en memoria
  - 4.2. Motores de data streaming

## 6. Cronograma

### 6.1. Cronograma de la asignatura \*

Sem	Actividad presencial en aula	Actividad presencial en laboratorio	Tele-enseñanza	Actividades de evaluación
1	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
2	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
3	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
4	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
5	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
6	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		<b>Práctica 1</b> TG: Técnica del tipo Trabajo en Grupo Evaluación continua Presencial Duración: 00:00
7	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
8	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
9	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
10	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		

11	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
12	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
13	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
14	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
15	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
16	<b>Clases de teoría</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral	<b>Práctica de laboratorio</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		<b>Práctica 2</b> TG: Técnica del tipo Trabajo en Grupo Evaluación continua Presencial Duración: 00:00
17				<b>Examen asignatura</b> EX: Técnica del tipo Examen Escrito Evaluación continua Presencial Duración: 03:00

Para el cálculo de los valores totales, se estima que por cada crédito ECTS el alumno dedicará dependiendo del plan de estudios, entre 26 y 27 horas de trabajo presencial y no presencial.

\* El cronograma sigue una planificación teórica de la asignatura y puede sufrir modificaciones durante el curso derivadas de la situación creada por la COVID-19.

## 7. Actividades y criterios de evaluación

### 7.1. Actividades de evaluación de la asignatura

#### 7.1.1. Evaluación continua

Sem.	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
6	Práctica 1	TG: Técnica del tipo Trabajo en Grupo	Presencial	00:00	20%	/ 10	CE07 CB04 CE05 CB03
16	Práctica 2	TG: Técnica del tipo Trabajo en Grupo	Presencial	00:00	20%	/ 10	CE07 CB04 CE05 CB03
17	Examen asignatura	EX: Técnica del tipo Examen Escrito	Presencial	03:00	60%	4 / 10	CE07 CB04 CE05 CB03

#### 7.1.2. Evaluación sólo prueba final

No se ha definido la evaluación sólo por prueba final.

#### 7.1.3. Evaluación convocatoria extraordinaria

Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
Examen convocatoria extraordinaria	EX: Técnica del tipo Examen Escrito	Presencial	03:00	60%	4 / 10	CE05 CB03 CE07 CB04

## 7.2. Criterios de evaluación

### Criterios para aprobar la asignatura por evaluación progresiva:

La asignatura se evaluará mediante 2 proyectos y un examen, siendo los pesos de estas partes los que se indican a continuación:

- Proyecto/entrega 1: 20%
- Proyecto/entrega 2: 20%
- Examen: 60%

Para aprobar la asignatura es necesario:

- Obtener en la calificación final un 5.
- Obtener, al menos, 4 puntos sobre 10 en el examen.

### Criterios para aprobar la asignatura en la convocatoria extraordinaria:

La evaluación en la convocatoria extraordinaria de la asignatura será mediante un examen final en el que se evaluarán todos los conocimientos del temario de la asignatura y donde será necesario sacar al menos 4 puntos sobre 10. El examen se realizará en la fecha establecida en el calendario oficial de exámenes por jefatura de estudios para la convocatoria correspondiente. Para calcular la nota final en esta convocatoria se tendrá en cuenta única y exclusivamente la nota del examen, que representa el 60% de la asignatura.

- Examen: 60%

### Nota final sobre las prácticas

Las prácticas 1 y 2 no son recuperables, pues se realizarán en base a un criterio de realización y evaluación continua durante las clases como si fueran prácticas de laboratorio.

## 8. Recursos didácticos

### 8.1. Recursos didácticos de la asignatura

Nombre	Tipo	Observaciones
Data Governance: How to Design, Deploy, and Sustain an Effective Data Governance Program (English Edition) 2o Edición. John Ladley	Bibliografía	
Data Provenance A Complete Guide - 2020 Edition (English Edition). Gerardus Blokdyk	Bibliografía	
Pandas 1.x Cookbook: Practical recipes for scientific computing, time series analysis, and exploratory data analysis using Python, 2nd Edition (English Edition). Matt Harrison; Theodore Petrou	Bibliografía	
SQL & NoSQL Databases: A Comprehensive Introduction To Relational (SQL) And Non-relational (NoSQL) Databases (English Edition). Danniella Hardin	Bibliografía	
SQL & NoSQL Databases: Models, Languages, Consistency Options and Architectures for Big Data Management (English Edition). Andreas Meier; Michael Kaufmann	Bibliografía	

Stream Processing with Apache Flink: Fundamentals, Implementation, and Operation of Streaming Applications. Fabian Hueske, Vasiliki Kalavri	Bibliografía	
---	--------------	--

## 9. Otra información

---

### 9.1. Otra información sobre la asignatura

Se prevé que la situación sanitaria causada por la pandemia de la COVID-19 haya mejorado lo suficiente como para permitir utilizar el aforo completo de las aulas. Por ello se ha planificado la docencia de este semestre en modo presencial.

Si las condiciones sanitarias no permitieran usar el aforo completo de las aulas, se pasará a una modalidad de presencialidad mixta por turnos, como el planteado en asignaturas de primer semestre, sin necesidad de modificar esta guía.