



1. Identificación

1.1. De la Asignatura

Curso Académico	2021/2022
Titulación	MÁSTER UNIVERSITARIO EN TECNOLOGÍAS DE ANÁLISIS DE DATOS MASIVOS: BIG DATA
Nombre de la Asignatura	BASES DE DATOS A GRAN ESCALA
Código	6418
Curso	PRIMERO
Carácter	OBLIGATORIA
N.º Grupos	1
Créditos ECTS	4.5
Estimación del volumen de trabajo del alumno	112.5
Organización Temporal/Temporalidad	Cuatrimestre
Idiomas en que se imparte	ESPAÑOL
Tipo de Enseñanza	Presencial

1.2. Del profesorado: Equipo Docente

Coordinación de la asignatura DIEGO SEVILLA RUIZ Grupo de Docencia: 1	Área/Departamento	ARQUITECTURA Y TECNOLOGÍA DE COMPUTADORES/ INGENIERÍA Y TECNOLOGÍA DE COMPUTADORES			
	Categoría	PROFESORES TITULARES DE UNIVERSIDAD			
	Correo Electrónico / Página web / Tutoría electrónica	dsevilla@um.es http://ditec.um.es/personal/1 Tutoría Electrónica: SÍ			
Coordinación de los grupos: 1	Teléfono, Horario y Lugar de atención al alumnado	Duración	Día	Horario	Lugar
		Anual	Miércoles	11:00- 14:00	868887571, Facultad de Informática B1.3.029



2. Presentación

El objetivo de esta materia es proporcionar al alumno los fundamentos para el modelado de datos siguiendo los paradigmas más utilizados dentro del entorno Big Data: SQL y NoSQL. Se estudiarán las diferencias de organización de los datos y su acceso, así como los diferentes paradigmas de organización de datos (clave/valor, columnas, grafos, etc). El alumno aprenderá a organizar los datos de tal forma que permita un acceso distribuido eficiente, tolerante a fallos y de alta disponibilidad. Finalmente, se iniciará al alumno en diferentes lenguajes de acceso a datos y a la exportación de los mismos en diferentes formatos y tecnologías.

3. Condiciones de acceso a la asignatura

3.1 Incompatibilidades

No consta

3.2 Recomendaciones

Se recomienda conocimiento del estándar SQL, bases de datos relacionales y lenguajes de programación Python.

4. Competencias

4.1 Competencias Básicas

- CB6. Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación
- CB7. Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio
- CB8. Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios
- 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.



4.2 Competencias de la titulación

- G1. Capacidad para afrontar tareas y situaciones críticas
- G2. Capacidad de trabajo autónomo y toma de decisiones
- G3. Capacidades asociadas al trabajo en equipo: cooperación, liderazgo, saber escuchar
- G4. Capacidad analítica, crítica y de síntesis
- G5.. Creatividad
- E16. Capacidad para seleccionar, atendiendo a criterios de eficiencia, escalabilidad, optimización de acceso, tolerancia a fallos y adecuación al entorno de producción, las bases de datos y el paradigma de datos óptimo en soluciones Big Data.

4.3 Competencias transversales y de materia

- Competencia 1. CM1. Comprender y ser capaz de usar eficazmente los paradigmas más extendidos de modelado de datos
- Competencia 2. CM2. Ser capaz de realizar modelado de datos relacional y usar efectivamente las herramientas asociadas
- Competencia 3. CM3. Ser capaz de realizar modelado de datos NoSQL y usar eficientemente las herramientas relacionadas

5. Contenidos

Bloque 1: Modelado de datos Relacional

TEMA 1. Introducción al modelado de datos

TEMA 2. Modelado relacional: SQL

Bloque 2: Modelado de datos NoSQL

TEMA 3. Introducción al modelado de datos NoSQL

TEMA 4. Bases de datos documentales

TEMA 5. Bases de datos columnares y manejo de series temporales

TEMA 6. Bases de datos de grafos

TEMA 7. Tecnologías NewSQL

PRÁCTICAS

Práctica 1. Modelado de datos relacional y SQL: Relacionada con los contenidos Bloque 1, Tema 1 y Tema 2

Práctica 2. BBDD NoSQL Documentales: Relacionada con los contenidos Tema 3 y Tema 4

Práctica 3. BBDD NoSQL Columnares y series temporales: Relacionada con los contenidos Tema 5

Práctica 4. BBDD NoSQL de Grafos: Relacionada con los contenidos Tema 6

Práctica 5. Práctica NewSQL: Relacionada con los contenidos Tema 7



6. Metodología Docente

Actividad Formativa	Metodología	Horas Presenciales	Trabajo Autónomo	Volumen de trabajo
Clases de pizarra	Se utilizarán clases magistrales, en las que se expone el contenido de cada tema. El alumno dispondrá de copias de los materiales formativos con anterioridad y el profesor promoverá una actitud activa, realizando preguntas que permitan aclarar aspectos concretos y dejando cuestiones abiertas para la reflexión del alumno.	13.5	33	46.5
Clases prácticas	Se realizarán tutoriales introductorios a las herramientas, donde los alumnos aplicarán los conceptos vistos en clase a la resolución de problemas sencillos. Los alumnos dispondrán también de los manuales en el caso de las herramientas utilizadas. Se realizará un trabajo tutelado en el que los alumnos tendrán que profundizar de manera autónoma pero tutelada en alguno de los contenidos básicos, o aplicar la metodología vista a la resolución de un problema real.	18	36	54
Tutorías y actividades de evaluación	Entrevista de prácticas, examen y defensa del trabajo final	1.5	10.5	12
	Total	33	79.5	112.5



Docencia en presencialidad adaptada

El Plan de Contingencia 4.0, aprobado por la Junta de Facultad el 11 de junio de 2021, tiene como objetivo primordial garantizar una presencialidad segura durante el curso 2021/22 que permita al estudiantado asistir al mayor número posible de clases. Para lograr este objetivo se articulará una rotación racional y equitativa de grupos presenciales. El Decanato será el responsable de realizar esta división de grupos rotatorios a principio de cuatrimestre según lo dispuesto en dicho Plan de Contingencia. La metodología en presencialidad adaptada a 1 m, o en su caso la que indiquen las autoridades sanitarias y académicas, se realizará de acuerdo con lo dispuesto en dicho plan.

7. Horario de la asignatura

<http://www.um.es/informatica/index.php?pagina=planificacion&subseccion=horarios>

8. Sistema de Evaluación

Métodos / Instrumentos	Evaluación de Pruebas periódicas y/o Prueba final
Criterios de Valoración	El profesor, a través de preguntas, o bien a través de un trabajo de envergadura final, evaluará el grado de cumplimiento de los objetivos de la asignatura por parte del alumno.
Ponderación	40



Métodos / Instrumentos	Evaluación de prácticas realizadas
Criterios de Valoración	El profesor evaluará la corrección de las prácticas, el grado de cumplimiento por parte del alumno así como la defensa de las mismas.
Ponderación	40
Métodos / Instrumentos	Seguimiento continuado y objetivable de una participación activa
Criterios de Valoración	
Ponderación	0
Métodos / Instrumentos	Evaluación de trabajos académicamente dirigidos
Criterios de Valoración	Se valorará el trabajo del alumno, el grado de conocimiento de la tecnología o materia tratada así como la defensa a las preguntas del profesor sobre el mismo.
Ponderación	20

Fechas de exámenes

<http://www.um.es/informatica/index.php?pagina=planificacion&subseccion=examenes>

9. Resultados del Aprendizaje

Los principales resultados del aprendizaje en esta materia, y su relación con las competencias de la titulación, son:

Resultados de aprendizaje	Relación con las competencias de la titulación		
	básicas	transversales y generales	específicas
Comprender y saber aplicar las diferentes formas de modelado de datos	CB6, CB8	T5 G1, G2, G4	E16
Conocer la dinámica de funcionamiento y	CB6, CB10	T5 G4	E16



las diferencias de los paradigmas SQL y NoSQL con sus diferentes variantes			
Ser capaz de diseñar un modelo de datos eficiente, distribuido y tolerante a fallos	CB6, CB7, CB8, CB9, CB10	T2, T5 G1, G3, G4	E16

10. Bibliografía

Bibliografía Básica



Sadalage, Fowler. *NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence*, Addison-Wesley, 2012



Redmond, E., Wilson, J. R. *Seven Databases in Seven Weeks: A Guide to Modern Databases and the NoSQL Movement*, Pragmatic Bookshelf, 2012



Marz, N., Warren, J. *Big Data: Principles and best practices of scalable realtime data systems*, Manning Publications, 2015

Bibliografía Complementaria



Tom White. *Hadoop: The Definitive Guide (4ed.)*, O'Reilly, 2015.



Rick Copeland. *MongoDB Applied Design Patterns*, O'Reilly, 2015.

11. Observaciones y recomendaciones

En las convocatorias de junio/julio los alumnos tendrán que hacer todas las entregas de prácticas especificadas a lo largo del curso, y habrá una entrevista con el profesor donde cada alumno deberá demostrar su conocimiento de la materia. El hecho de no presentar las prácticas o el examen final/trabajo final hará que el alumno aparezca como "NO PRESENTADO". El alumno aprobará si la suma de los trabajos presentados es superior o igual a 5 puntos, teniendo en cuenta la ponderación.



Para los casos de realización fraudulenta de ejercicios o pruebas será de aplicación el recogido en la Normativa de evaluación del rendimiento académico de los estudiantes y de revisión de calificaciones.

NECESIDADES EDUCATIVAS ESPECIALES. Aquellos estudiantes con discapacidad o necesidades educativas especiales podrán dirigirse al Servicio de Atención a la Diversidad y Voluntariado (ADYV; <http://www.um.es/advv/>) para recibir orientación sobre un mejor aprovechamiento de su proceso formativo y, en su caso, la adopción de medidas de equiparación y de mejora para la inclusión, en virtud de la Resolución Rectoral R-358/2016. El tratamiento de la información sobre este alumnado, en cumplimiento con la LOPD, es de estricta confidencialidad.