



## FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
Nombre completo	Datos Masivos / Big Data
Código	DTC-BA-423
Título	<a href="#">Grado en Análisis de Negocios / Business Analytics por la Universidad Pontificia Comillas</a>
Impartido en	Grado en Ingeniería en Tecnologías de Telecom. y Grado en Análisis de Negocios/Business Analytics [Cuarto Curso]
Nivel	Reglada Grado Europeo
Cuatrimestre	Semestral
Créditos	3,0 ECTS
Carácter	Obligatoria (Grado)
Departamento / Área	Departamento de Telemática y Computación
Responsable	Moisés Martínez
Horario	Consultar horarios de la escuela
Horario de tutorías	Concertar cita con el profesor

### Datos del profesorado

## DATOS ESPECÍFICOS DE LA ASIGNATURA

### Contextualización de la asignatura

#### Aportación al perfil profesional de la titulación

La tecnología Big Data es esencial en la disciplina de Analítica avanzada y procesado masivo de datos y una herramienta de uso cada vez más extendido en las empresas como habilitador de los análisis avanzados. De esta manera el objetivo de la asignatura es que el alumno aprenda los conceptos relacionados de Big data, desde sus origen e historia, características y diferencias respecto a sistemas tradicionales, su arquitectura paralela y las diferencias con otros sistemas paralelos y ecosistema del ciclo del dato básico, desde su adquisición del dato, almacenamiento al procesamiento. Se relacionará las capacidades que ofrece esta tecnología con los problemas que para los que esta especialmente indicada, las necesidades que cubre y los casos de uso que consiguen aportar valor diferencial a las empresas y la sociedad.

Al finalizar el curso, los alumnos entenderán los bloques habituales de una instalación tipo, el funcionamiento de big data, su utilidad y sabrán en qué casos debe ser aplicado, por qué y para qué, de tal manera que puedan proponer y defender casos concretos de uso.

#### Prerequisitos



Ninguno

## Competencias - Objetivos

### Competencias

#### GENERALES

<b>CG01</b>	Capacidad de organización y planificación en la identificación de problemas en el contexto de datos masivos	
	<b>RA1</b>	Identificar y organizar temporalmente las tareas necesarias para la realización sus actividades de aprendizaje, cumpliendo con los plazos establecidos de entrega de actividades.
	<b>RA2</b>	Ser capaz de integrarse y participar en el desarrollo organizado de un trabajo en grupo.
<b>CG02</b>	Capacidad de análisis de datos masivos procedentes de diversas fuentes: texto, audio, numérica e imagen	
	<b>RA1</b>	Describir, relacionar e interpretar situaciones y planteamientos sencillos
	<b>RA2</b>	Seleccionar las fuentes de información digital más significativas
	<b>RA3</b>	Identificar las fuentes de información más relevantes para cada problema
	<b>RA4</b>	Utilizar las herramientas digitales idóneas para la adquisición y gestión de la información y es capaz de integrar esa información de una manera racional
<b>CG03</b>	Resolución de problemas y toma de decisiones en un entorno de datos masivos tanto cuantitativos como cualitativos	
	<b>RA1</b>	Resolver problemas y realizar trabajos aplicando una diversidad metodológica de análisis
	<b>RA2</b>	Identificar correctamente los conocimientos aplicables a cada situación
	<b>RA3</b>	Identificar la metodología más apropiada para la resolución de los problemas planteados, conociendo las herramientas software más relevantes y aportando una visión crítica.
<b>CG07</b>	Capacidad de liderazgo y trabajo en equipo, en la sociedad de la información	
	<b>RA1</b>	Participar de forma activa en el trabajo de grupo compartiendo información, conocimientos y experiencias



	<b>RA2</b>	Orientarse a la consecución de acuerdos y objetivos comunes
	<b>RA3</b>	Ser capaz de organizar y/o coordinar equipos de trabajo
<b>CG08</b>	Capacidad crítica y autocrítica en la sociedad de la información	
	<b>RA1</b>	Identificar los supuestos y las limitaciones de los modelos y metodologías a aplicar en cada situación
	<b>RA2</b>	Ejercer una valoración crítica de los modelos y metodologías a aplicar en cada situación
<b>CG09</b>	Compromiso ético en la sociedad de la información	
	<b>RA1</b>	Ser honesto en el desarrollo de las actividades académicas y en otros aspectos de la vida y no ser pasivo ante la deshonestidad de otros.
	<b>RA2</b>	Conocer y desarrollar las bases de la ética profesional
<b>CG11</b>	Capacidad para aprender y trabajar autónomamente en la sociedad de la información	
	<b>RA1</b>	Organizar los materiales de la asignatura con el fin de articular su aprendizaje
	<b>RA2</b>	Poseer criterio para decidir la utilización de los medios humanos y otros recursos puestos a su disposición para el aprendizaje
	<b>RA3</b>	Realizar sus trabajos y su actividad necesitando sólo unas indicaciones iniciales y un seguimiento básico
<b>ESPECÍFICAS</b>		
<b>CE25</b>	Comprender los conceptos más importantes asociados al Big Data y sus elementos característicos	
	<b>RA1</b>	Conocer el concepto de paralelismo y la importancia de la distribución de datos voluminosos
	<b>RA2</b>	Identificar las necesidades específicas asociadas al manejo, adquisición y almacenamiento de volúmenes masivos de datos.
	<b>RA3</b>	Comprender y conocer las tecnologías de comunicaciones (redes) y computación precisas para la distribución y recuperación de información, así como de su análisis.
<b>CEO34</b>	Entender la computación y las bases de datos distribuidas, con énfasis en la arquitectura Hadoop y las herramientas actuales más extendidas en el contexto de Big Data.	



<b>RA01</b>	Conoce y comprende la problemática del Big Data y las técnicas más avanzadas de computación y almacenamiento usadas en la actualidad.
<b>RA02</b>	Comprende y es capaz de administrar una arquitectura Hadoop y otras arquitecturas NoSQL.
<b>RA03</b>	Comprende y aplica las técnicas de MapReduce y otras técnicas avanzadas de computación específicas para el análisis de datos voluminosos.

## BLOQUES TEMÁTICOS Y CONTENIDOS

### Contenidos – Bloques Temáticos

#### 1: Introducción a Big Data

1. ¿Qué es Big Data y qué no es Big Data?
2. El valor del dato
3. Volúmenes y fuentes de datos. Información estructurada y no estructurada.
4. Ciclo de vida de un proyecto Big Data. Perfiles profesionales.

#### 2: Sistemas Distribuidos

1. Introducción.
2. Problemas. Tolerancia a fallos. Balanceo. Disponibilidad. Redundancia.
3. Sistemas operativos basados en Linux.
4. Virtualización:
  - Introducción y plataformas.
  - Máquinas virtuales vs contenedores

#### 3: Ecosistema Hadoop

1. Introducción a Hadoop. Componentes.
2. Arquitectura Hardware y Software.
3. Administración y monitorización de un cluster.

#### 4: Almacenamiento Masivo

1. Sistema de ficheros HDFS.
2. Servicios y roles.
3. HUE.

#### 5: Introducción al Procesamiento Masivo

1. YARN
2. MapReduce
3. Spark

#### 6: Cloud & Big Data



1. Infraestructura on-premise vs Cloud.
2. Concepto "as a service".
3. Proveedores y plataformas Cloud.

### 7: Tecnología Big Data en la empresa

1. Impacto de Big Data en Analytics y la Inteligencia Artificial/Machine Learning
2. Big Data como tecnología exponencial y disruptiva en el negocio.
3. Casos de uso de tecnología Big Data.

## METODOLOGÍA DOCENTE

### Aspectos metodológicos generales de la asignatura

#### Metodología Presencial: Actividades

1. **Clase magistral y presentaciones generales.** Exposición de los principales conceptos y procedimientos mediante la explicación por parte del profesor. Incluirá presentaciones dinámicas, pequeños ejemplos prácticos y la participación reglada o espontánea de los estudiantes. Además, el profesor propondrá la lectura de temas preparados por el profesor, artículos científicos y divulgativos que permitirán introducir o profundizar en un tema, con el objetivo de aportar ideas y generar debate **(24 horas). CE 25,CG 03, CG 02 CG 01, CB 03, CB 02**
2. **Resolución en clase de problemas prácticos.** Resolución de unos primeros problemas para situar al alumno en contexto. La resolución correrá a cargo del profesor y los alumnos de forma cooperativa **(2 horas). CE 25,CG 03, CG 02 CG 01, CB 03, CB 02**
3. **Prácticas.** Cada alumno realizará de forma aislada o en grupo una serie de prácticas regladas. Las prácticas finalizarán con la redacción de un informe o la redacción de las distintas experiencias **(4 horas). CE 25,CG 03, CG 02 CG 01, CB 03, CB 02**

#### Metodología No presencial: Actividades

1. **Estudio individual del material.** Actividad realizada individualmente por el estudiante para preparar el tema que se discutirá en clase. **(60 horas). CE 25,CG 03, CG 02 CG 01, CB 03, CB 02**
2. **Resolución de problemas prácticos** a resolver fuera del horario de clase por parte del alumno. El alumno debe utilizar e interiorizar los conocimientos aportados en la materia. La corrección a la clase se realizará por parte de alguno de los alumnos o el profesor según los casos. La corrección individualizada de cada ejercicio la realizará el propio alumno u otro compañero según los casos (método de intercambio) **(10 horas).CE 25,CG 03, CG 02 CG 01, CB 03, CB 02**
3. **Prácticas de la asignatura.** Realización de dos prácticas que plasme los conocimientos adquiridos en la asignatura. **(10 horas).CE 25,CG 03, CG 02 CG 01, CB 03, CB 02**

## RESUMEN HORAS DE TRABAJO DEL ALUMNO

HORAS PRESENCIALES		
Lecciones de Carácter expositivo	Exposición pública de temas o trabajos	Ejercicios y resolución de casos y de problemas



24.00	2.00	6.00
<b>HORAS NO PRESENCIALES</b>		
Estudios individual y/o en grupo, y lectura organizada	Ejercicios y resolución de casos y de problemas	Trabajos monográficos y de investigación, individuales o colectivos
30.00	18.00	10.00
<b>CRÉDITOS ECTS: 3,0 (90,00 horas)</b>		

## EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

Actividades de evaluación	Criterios de evaluación	Peso
<b><u>Realización de exámenes:</u></b>	<ul style="list-style-type: none"> <li>Examen final.</li> <li>Comprensión de conceptos mediante pruebas abiertas, enfocado a casos prácticos.</li> <li>Aplicación de los conceptos adquiridos</li> <li>Presentación y comunicación escrita.</li> </ul>	60 %
<b><u>Evaluación del rendimiento teórico continuo:</u></b>	<ul style="list-style-type: none"> <li>Realización de pruebas de seguimiento, debates y ejercicios.</li> <li>Comprensión de conceptos mediante pruebas objetivas y one minute papers.</li> <li>Debate y análisis de los documentos técnico/científicos.</li> <li>Análisis e interpretación de los resultados obtenidos en la resolución de problemas mediante debates</li> </ul>	30 %
<b><u>Evaluación del rendimiento práctico continuo:</u></b>	<ul style="list-style-type: none"> <li>Casos a desarrollar en la clase de prácticas</li> <li>Comprensión de conceptos.</li> <li>Aplicación de conceptos al desarrollo de casos sobre los nuevos conceptos planteados.</li> <li>Aplicación de todos los conocimientos adquiridos durante el curso.</li> </ul>	10 %



## Calificaciones

### Criterios de Calificación

La calificación en la **convocatoria ordinaria** de la asignatura se obtendrá como:

- Un 60% la calificación del examen final.
- Un 30% será la calificación de las pruebas prácticas
  - 10% corresponde a prácticas grupales
  - 10% prácticas individuales
  - 10% breves test- exámenes- entregas de prácticas guiadas realizadas en clase.
- Un 10% será la calificación de la asistencia y participación en clase.

La calificación en la **convocatoria extraordinaria**:

- Un 60% la calificación del examen final.
- Un 30% será la calificación de las pruebas prácticas
  - 10% corresponde a prácticas grupales
  - 10% prácticas individuales
  - 10% breves test- exámenes- entregas de prácticas guiadas realizadas en clase.
- Un 10% será la calificación de la asistencia y participación en clase.

**Para aprobar la asignatura los alumnos tienen que tener al menos 5 puntos sobre 10 en el examen final, y en las práctica, tanto en la convocatoria ordinaria como en la extraordinaria.**

## PLAN DE TRABAJO Y CRONOGRAMA

Actividades	Fecha de realización	Fecha de entrega
<p><b>Actividades Presenciales y No presenciales</b></p> <p><b>Fecha de realización</b></p> <p><b>Fecha de entrega</b></p> <ul style="list-style-type: none"><li>• Lectura y estudio de los contenidos teóricos Después de cada clase</li><li>• Resolución de los problemas propuestos o debates</li></ul>		



Semanalmente		
<ul style="list-style-type: none"><li>Preparación de las prácticas</li></ul>		
Semanas		
8, 9 y 10		
Semanas		
11		
<ul style="list-style-type: none"><li>Preparación del examen final</li></ul>		

## BIBLIOGRAFÍA Y RECURSOS

### Bibliografía Básica

Notas preparadas por el profesor disponibles en Moodle.

### Bibliografía Complementaria

- Big Data: A Revolution That Will Transform How We Live, Work, and Think. [Viktor Mayer-Schonberger](#) (Autor), [Kenneth Cukier](#) (Autor)
- George Orwell. 1984 . ISBN: 9788499890944
- Michael Lewis. Moneyball: The Art of Winning an Unfair Game Paperback – March 17, 2004. ISBN: 978-0393324815 or the film 'Moneyball (2011)'. Sony Pictures Director: Bennett Miller

En cumplimiento de la normativa vigente en materia de **protección de datos de carácter personal**, le informamos y recordamos que puede consultar los aspectos relativos a privacidad y protección de datos que ha aceptado en su matrícula entrando en esta web y pulsando "descargar"

<https://servicios.upcomillas.es/sedelectronica/inicio.aspx?csv=02E4557CAA66F4A81663AD10CED66792>