



CAMPUS  
DE EXCELENCIA  
INTERNACIONAL

PROCESO DE  
COORDINACIÓN DE LAS  
ENSEÑANZAS PR/CL/001



E.T.S. de Ingenieros  
Industriales

# ANX-PR/CL/001-01

## GUÍA DE APRENDIZAJE

### ASIGNATURA

**53001521 - Tools for big data analytics**

### PLAN DE ESTUDIOS

05BD - Master Universitario en Ingeniería de la Organización

### CURSO ACADÉMICO Y SEMESTRE

2018/19 - Primer 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

## 1. Datos descriptivos

---

### 1.1. Datos de la asignatura

<b>Nombre de la asignatura</b>	53001521 - Tools for big data analytics
<b>No de créditos</b>	3 ECTS
<b>Carácter</b>	Optativa
<b>Curso</b>	Segundo curso
<b>Semestre</b>	Tercer semestre
<b>Período de impartición</b>	Septiembre-Enero
<b>Idioma de impartición</b>	Castellano
<b>Titulación</b>	05BD - Master universitario en ingeniería de la organización
<b>Centro en el que se imparte</b>	05 - Escuela Técnica Superior de Ingenieros Industriales
<b>Curso académico</b>	2018-19

## 2. Profesorado

---

### 2.1. Profesorado implicado en la docencia

<b>Nombre</b>	<b>Despacho</b>	<b>Correo electrónico</b>	<b>Horario de tutorías</b> *
Joaquin Bienvenido Ordieres Mere (Coordinador/a)	UD Proyectos	j.ordieres@upm.es	X - 08:30 - 10:30 X - 12:30 - 15:30 V - 13:30 - 15:30

\* 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

- Bigdata: de los datos a las decisiones
- Tecnología, organización y nuevos modelos de negocio digital
- Métodos cuantitativos avanzados para la gestión

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

- SO Linux a nivel usuario
- Programación en Python

## 4. Competencias y resultados de aprendizaje

---

### 4.1. Competencias

CB06 - 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

CE01 - Conocer y aplicar técnicas y herramientas para el manejo y análisis de grandes masas de datos

CE02 - Conocer y aplicar técnicas cuantitativas para la modelización y resolución de problemas de las organizaciones

CT03 - Diseña. Habilidad para diseñar un sistema, componente o proceso que alcance los requisitos deseados teniendo en cuenta restricciones realistas tales como las económicas, medioambientales, sociales, políticas, éticas, de salud y seguridad, de fabricación y de sostenibilidad

## 4.2. Resultados del aprendizaje

RA12 - Comprender y evaluar los efectos de la incorporación de las Tecnologías de la Información y las Comunicaciones (TIC) en todos los niveles de la organización

RA1 - Elegir y aplicar técnicas de análisis exploratorios de grandes bases de datos

RA13 - Identificar y analizar el impacto que las tecnologías digitales, de la comunicación y los datos y de la inteligencia artificial tienen en la aparición de nuevos modelos de negocio disruptivos

## 5. Descripción de la asignatura y temario

---

### 5.1. Descripción de la asignatura

La asignatura que se imparte en el tercer semestre del Master en Ingeniería de Organización, con carácter optativo, tiene como objetivos fundamentales que los alumnos sean capaces de:

- Comprender el significado de las soluciones BigData, algunos casos de uso y valorar la diferencia de paradigma con las soluciones IT convencionales en la empresa
- Conocer y aplicar las herramientas que les permitan utilizar técnicas ETL en esos entornos.
- Aplicar estas herramientas y metodologías para la explotación de datos a través de modelos.
  
- Aplicar estas herramientas y metodologías para la visualización de datos en esos entornos.

La asignatura tiene un carácter eminentemente práctico

Se llevarán a cabo las actividades en un cluster de amazon en clase, ecludas aquellas sesiones introductorias. El alumno deberá llevar a cabo las actividades también fuera de clase.

Dispondrá de de un foro para discutir los comandos que no le funcionen y compartir las experiencias.

## 5.2. Temario de la asignatura

1. INTRODUCCIÓN Y FUNDAMENTOS ECONÓMICOS
  - 1.1. ¿Por qué el Big Data es necesario? ¿quién lo utiliza en su actividad diaria?
  - 1.2. El IoT. Uso de BigData en Business
  - 1.3. Almacenamiento de datos en BigData. Introducción a la Tecnología
2. ARQUITECTURA BIG-DATA
  - 2.1. Clusters
  - 2.2. Bases de Datos en BigData
  - 2.3. Desplegando un entorno cluster Hadoop
3. BIGDATA ANALYTICS
  - 3.1. Operaciones con los datos almacenados
  - 3.2. El proyecto SPARK
  - 3.3. Desplegando un entorno cluster Spark
4. SOLUCIONES COMERCIALES Y SOLUCIONES HOME-MADE
  - 4.1. Grandes Proveedores Horizontales
  - 4.2. El mundo OpenSource
5. EL PROYECTO PRACTICO DE GRUPO
  - 5.1. Elaboración de un proyecto por grupos
  - 5.2. ETL de los datos
  - 5.3. Modelización
  - 5.4. Presentación de los resultados

## 6. Cronograma

### 6.1. Cronograma de la asignatura \*

Sem	Actividad presencial en aula	Actividad presencial en laboratorio	Otra actividad presencial	Actividades de evaluación
1	<b>Introducción Desarrollar expositivamente los contenidos del tema 1</b> Duración: 02:00 LM: Actividad del tipo Lección Magistral			<b>Gamificación para valorar conocimientos de la última lección vista</b> ET: Técnica del tipo Prueba Telemática Evaluación continua Duración: 00:05
2	<b>Introducción Desarrollar expositivamente los contenidos del tema 1</b> Duración: 01:00 LM: Actividad del tipo Lección Magistral			<b>Gamificación para valorar conocimientos de la última lección vista</b> ET: Técnica del tipo Prueba Telemática Evaluación continua Duración: 00:05
3	<b>Presentación sobre arquitecturas</b> Duración: 01:00 LM: Actividad del tipo Lección Magistral	<b>Aplicación de los conocimientos vistos en teoría de un modo práctico</b> Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio		<b>Gamificación para valorar conocimientos de la última lección vista</b> ET: Técnica del tipo Prueba Telemática Evaluación continua Duración: 00:05
4	<b>Presentación sobre Big Analytics</b> Duración: 01:00 LM: Actividad del tipo Lección Magistral	<b>Aplicación de los conocimientos vistos en teoría de un modo práctico</b> Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio		<b>Gamificación para valorar conocimientos de la última lección vista</b> ET: Técnica del tipo Prueba Telemática Evaluación continua Duración: 00:05
5	<b>Valoración de diferentes alternativas de implementación</b> Duración: 01:00 LM: Actividad del tipo Lección Magistral	<b>Aplicación de los conocimientos vistos en teoría de un modo práctico</b> Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio		<b>Gamificación para valorar conocimientos de la última lección vista</b> ET: Técnica del tipo Prueba Telemática Evaluación continua Duración: 00:05
6	<b>Definición del trabajo Grupal</b> Duración: 01:00 AC: Actividad del tipo Acciones Cooperativas	<b>Aplicación de los conocimientos vistos en teoría de un modo práctico</b> Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio		<b>Gamificación para valorar conocimientos de la última lección vista</b> ET: Técnica del tipo Prueba Telemática Evaluación continua Duración: 00:05
7		<b>Aplicación de los conocimientos vistos en teoría de un modo práctico</b> Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio	<b>Desarrollo del trabajo práctico de grupo</b> Duración: 01:00 OT: Otras actividades formativas	<b>Gamificación para valorar conocimientos de la última lección vista</b> ET: Técnica del tipo Prueba Telemática Evaluación continua Duración: 00:05
8		<b>Aplicación de los conocimientos vistos en teoría de un modo práctico</b> Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio	<b>Desarrollo del trabajo práctico de grupo</b> Duración: 01:00 OT: Otras actividades formativas	<b>Gamificación para valorar conocimientos de la última lección vista</b> ET: Técnica del tipo Prueba Telemática Evaluación continua Duración: 00:05
9		<b>Aplicación de los conocimientos vistos en teoría de un modo práctico</b> Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio	<b>Desarrollo del trabajo práctico de grupo</b> Duración: 01:00 OT: Otras actividades formativas	

10		<b>Aplicación de los conocimientos vistos en teoría de un modo práctico</b> Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio	<b>Desarrollo del trabajo práctico de grupo</b> Duración: 01:00 OT: Otras actividades formativas	
11		<b>Aplicación del trabajo de grupo y desarrollo de soluciones</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
12		<b>Aplicación del trabajo de grupo y desarrollo de soluciones</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
13		<b>Aplicación del trabajo de grupo y desarrollo de soluciones</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		
14		<b>Aplicación del trabajo de grupo y desarrollo de soluciones</b> Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio		<b>Entrega de trabajo individual</b> TI: Técnica del tipo Trabajo Individual Evaluación sólo prueba final Duración: 00:30
15				<b>Presentación de resultados</b> PG: Técnica del tipo Presentación en Grupo Evaluación continua Duración: 02:00  <b>Entrega de vídeo justificativo</b> PI: Técnica del tipo Presentación Individual Evaluación sólo prueba final Duración: 00:30
16				

Las horas de actividades formativas no presenciales son aquellas que el estudiante debe dedicar al estudio o al trabajo personal.

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.

## 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
1	Gamificación para valorar conocimientos de la última lección vista	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	1%	3 / 10	CE01 CE02
2	Gamificación para valorar conocimientos de la última lección vista	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	1%	3 / 10	CE01 CE02
3	Gamificación para valorar conocimientos de la última lección vista	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	1%	3 / 10	CE01 CE02
4	Gamificación para valorar conocimientos de la última lección vista	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	1%	3 / 10	CE01 CE02
5	Gamificación para valorar conocimientos de la última lección vista	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	1%	3 / 10	CE01 CE02
6	Gamificación para valorar conocimientos de la última lección vista	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	1%	3 / 10	CE01 CE02
7	Gamificación para valorar conocimientos de la última lección vista	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	1%	3 / 10	CE01 CE02
8	Gamificación para valorar conocimientos de la última lección vista	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	1%	3 / 10	CE01 CE02

15	Presentación de resultados	PG: Técnica del tipo Presentación en Grupo	Presencial	02:00	92%	5 / 10	CE01 CE02 CB06 CT03
----	----------------------------	--	------------	-------	-----	--------	------------------------------

### 7.1.2. Evaluación sólo prueba final

Sem	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
14	Entrega de trabajo individual	TI: Técnica del tipo Trabajo Individual	No Presencial	00:30	75%	5 / 10	CT03 CB06 CE01 CE02
15	Entrega de vídeo justificativo	PI: Técnica del tipo Presentación Individual	No Presencial	00:30	25%	5 / 10	CB06 CT03

### 7.1.3. Evaluación convocatoria extraordinaria

No se ha definido la evaluación extraordinaria.

## 7.2. Criterios de evaluación

### TRABAJOS EN GRUPO

La realización y presentación en clase de un trabajo en grupo es obligatoria, y la calificación de dicho trabajo formará parte de la calificación final de la asignatura para cada alumno. Estos trabajos tendrán como objetivo el abordar un problema y tratarlo en estos entornos, para poder comprender mejor el trabajo técnico diario que está implicado y ser conscientes de los pasos elementales.

El tema del trabajo podrá ser propuesto por el profesor o por el alumno, pero en este último caso siempre deberá contar con la aceptación previa del profesor.

Se procurará que las temáticas sean diversas y que la fuente de datos a utilizar sea pública. En algunos casos el profesor podrá proporcionar un scraper, si ello fuese de utilidad.

Los grupos de trabajo constarán de un mínimo de tres y un máximo de nueve alumnos.

Se empleará una técnica de evaluación interna de 360º

### SISTEMA DE CALIFICACIÓN

El sistema de calificación de la asignatura contempla dos modalidades alternativas a elegir por el alumno:

#### **Evaluación continua (Bolonia):**

- Una prueba de conocimientos gamificada.
- La calificación de los trabajos en grupo

La calificación final del curso en evaluación continua será el resultado de:

- Nota de la prueba de control, siempre que esté por encima de 3 puntos: 10%
- Trabajo en equipo: 80%
- Valoración del profesor (asistencia, participación, ?): 10%

Si no se aprueba por evaluación continua, el alumno deberá obligatoriamente presentarse al examen final.

#### Examen final (método tradicional):

- Prueba Práctica Individual (75%)
- Video explicativo de la actividad (25%)

## 8. Recursos didácticos

---

### 8.1. Recursos didácticos de la asignatura

Nombre	Tipo	Observaciones
Data-Intensive Text Processing with MapReduce by Jimmy Lin and Chris Dyer Morgan & Claypool Publishers, 2010.	Bibliografía	
Hadoop Real World Solutions Cookbook by Jonathan R. Owens, Brian Femiano, and Jon Lentz Publication Date: February 7, 2013   ISBN-10: 1849519129   ISBN-13:	Bibliografía	

978-1849519120		
BlueMix	Equipamiento	Plataforma de uso de IBM
Amazon	Equipamiento	Plataforma de desarrollo
GCP	Equipamiento	Plataforma de desarrollo