



# BIG DATA PROFESSIONAL CERTIFICATE (BDPC™)



## Introducción

Big Data es una de las tecnologías de la industria 4.0 de más alto impacto. Esta se define a partir de 7Vs que caracterizan los diferentes aspectos a considerar al manejar inmensos conjuntos de datos de fuentes variadas. Esta certificación está pensada para enseñar tanto a profesionales como a organizaciones los conceptos y evolución de Big Data; las diferencias entre data e información y conocimiento; la importancia de la gobernanza de los datos; las aplicaciones, metodologías y proyectos con Big Data; así como las tecnologías y frameworks que forman parte de Big Data.

## Objetivos

- Entender Big Data y sus extensas aplicaciones en el mercado
- Conocer las diferencias entre Big Data y data tradicional
- Aprender sobre las 7Vs del Big Data
- Conocer los diferentes frameworks que existen para Big Data
- Conceptualizar el ecosistema y las arquitecturas que se manejan con Big Data
- Relacionar casos de uso a Big Data

## Contenido

### Conceptos iniciales

Datos masivos  
Big Data  
Las 7Vs del Big Data  
MapReduce

### Datos, información y conocimiento

Los datos  
Las bases de datos  
Modelos de datos  
Procesamiento de datos

### Información

Características dependientes del proceso utilizado, el dominio y la interpretación  
Clústeres de datos o sistemas distribuidos

### Conocimiento

Perspectiva, eficacia y propósito  
La inteligencia artificial para generar conocimiento

### Evolución de Big Data

Las tres grandes fases de la evolución de Big Data  
Los puntos clave a través del tiempo para el Big Data  
Tendencias actuales

## **Gobernanza de datos**

- Definición
- Principales actividades del gobierno de datos
- Ventajas del gobierno de datos
- Roles en el gobierno de datos

## **Aplicaciones de Big Data**

### **Proyectos De Big Data con Business Intelligence (BI)**

- ¿Qué es el BI?
- Métodos de BI
- Los data warehouses (almacenes de datos)
- Evolución del BI y Analytics frente a orientación y técnicas

### **La metodología CRISP-DM. Big Data para la Minería de datos**

- Metodología CRISP-DM
- Metodología CRISP-DM como evolución de la metodología SEMMA

## **Análisis de datos**

- Tipos de análisis
- Aprendizaje supervisado
- Aprendizaje no supervisado
- Habilidades para el análisis de datos

## **Tecnologías para Big Data**

- Arquitectura de referencia NIST
- Arquitectura Big Data propuesta por Microsoft
- Arquitectura MapReduce
- Arquitectura NoSQL

## **Frameworks de datos masivos**

- Frameworks de Apache
- Otros frameworks de datos masivos
- Computación en la nube (Cloud computing)