



Big Data (Nivel Básico – Intermedio)

Objetivo General

El presente curso proporciona al perfil un conocimiento básico-intermedio de lo que es el mundo de Big Data, aportará conocimiento de programación en lenguaje **Scala** con el **Framework Spark** en integración con **Cludera Distribution Hadoop** (CDH) y la utilización de **Bases de Datos Operacionales**.

Este curso está enfocado a personas del mundo de sistemas que cuenten con los conocimientos básicos de programación y que tengan el deseo de seguir preparándose para aportar a su perfil herramientas y conocimientos de Big Data.

Requerimientos mínimos:

1. Deseables conocimientos de comandos Linux
2. Necesario saber programar (cualquier lenguaje).
3. Equipo de cómputo:
 - 8GB de RAM o superior.
 - 30GB disponibles de almacenamiento.

Curso: 40 horas

Distribución: 8 clases (sábados).

TEMARIO BIG DATA (Nivel Básico - Intermedio)

Módulo 1: (5 horas)

- I. Introducción al Mundo de BigData
 1. Ciencia de Datos
 2. Evolución de las Bases de Datos
 3. BigData

- II. Conociendo Apache Hadoop
 1. Instalación de Máquina Virtual (Cloudera)
 2. Configuración del ambiente en modo local

Módulo 2: (5 horas)

- III. Conociendo Apache Spark (1.6)
 1. Introducción de Spark.
 2. Ambiente de entrenamiento Spark.

- VI. Lenguaje de programación SCALA
 1. Introducción a Scala
 2. Aplicación de Scala en Spark

Módulo 3: (10 horas)

- V. Integración SCALA – Spark
 1. Introducción a SCALA-Spark
 - a. Transformaciones y acciones de RDDs simple
 - b. Transformaciones y acciones de RDDs clave/valor
 2. Lectura y escritura de Archivos (HDFS)
 3. Introducción Spark-SQL
 - a. Spark-SQL creación de DataFrames
 - b. Spark-SQL operaciones con DataFrames
 - c. Spark-SQL con Hive
 - d. Spark-SQL y bases de datos operacionales (HBASE / MONGODB)

Módulo 4: (5 hora)

- VI. Otros componentes de Apache Spark
 1. Spark Streaming
 2. Spark MLlib
 3. Spark GraphX

Módulo 5: (15 hora)

- VII. **Integración** de elementos para desarrollo de **ingestas, procesamiento, extracción y visualización de datos.**

[RESERVAR MI LUGAR \(clic\)](#)