

> CURSO BIG DATA FOR AZURE HDINSIGHT

DURACIÓN: **64 HORAS ACADÉMICAS**

MODALIDAD: **ONLINE**

En este curso, aprenderás a identificar y utilizar herramientas para la gestión de conjuntos de datos extremadamente grandes y complejos, que no se pueden procesar con herramientas tradicionales. HDInsight proporcionará el entorno de procesamiento y análisis distribuido permitiendo implementar soluciones empresariales.

I. DIRIGIDO A

- Profesionales, egresados y estudiantes, interesados en aprender de herramientas Big Data.

II. METODOLOGÍA

- El enfoque práctico-reflexivo de cada una de las sesiones virtuales en tiempo real, así como la participación de los estudiantes en un ambiente interactivo de clases moderado por el instructor, permitirá el aprendizaje de contenidos y su aplicación en la resolución de contextos reales, utilizando herramientas digitales para el aprendizaje significativo.

III. BENEFICIOS



- Acceso a la plataforma de Microsoft Azure HDInsight, desde cualquier PC o laptop con una conexión de Internet de 2 Mbps como mínimo
- La máquina virtual estará configurada con el software y hardware necesario, con velocidad 2.1 Gbps de navegación disponible para las sesiones de clases.
- Correo institucional de Cibertec.
- Office 365 Web (Word, Excel, Power Point, etc.)
- Acceso a Microsoft OneDrive y Microsoft Teams.

IV. LOGROS DEL CURSO



Al finalizar el curso, el alumno será capaz de diseñar una solución de inteligencia de Big Data para consumirse con herramientas de visualización especializadas. Adicionalmente, estará preparado para:

- Manipular la plataforma y herramientas de Azure HDInsight.
- Aprovechar la potencia de cómputo de clústeres de servidores eficientemente.
- Gestionar herramientas de visualización especializadas para soluciones Big Data.

V. CERTIFICACIONES



- Al aprobar la capacitación, obtendrás un certificado en Big Data for Azure HDInsight a nombre de Cibertec.

CERTIFICACIONES ASOCIADAS

- Especialización en Inteligencia de Negocios y Análítica de Datos
- Actualización en Inteligencia de Negocios

VI. PRERREQUISITOS



- Conocimientos de modelamiento de bases de datos relacionales (nivel intermedio)
- Conocimientos de expresiones de análisis de datos (nivel intermedio)
- Conocimientos de inteligencia de negocios (nivel intermedio)
- Conocimientos de inglés técnico (nivel intermedio)

CAPACITACIÓN RECOMENDADA

- Business Intelligence – SQL Server 2019
- Business Intelligence – SQL Server 2022
- SQL y Modelamiento de Datos – SQL Server 2019
- SQL y Modelamiento de Datos – SQL Server 2022
- Data Analytics – Power BI – Nivel Básico

VII. CONCEPTOS Y TERMINOLOGÍA

- Big Data
- Hadoop
- Spark
- Health Check
- HDFS
- Jobs

- Map Reduce
- Data Visualization
- Machine Learning
- Datawarehouse
- Cloud Computing
- Automatización

- Pig
- Data Lake
- Flume
- Sqoop
- Oozie
- Cluster

- Scala
- Hive
- Spark
- Python
- Business Analytics
- NoSQL

CONTENIDO TEMÁTICO

1 Fundamentos de Big Data

- Big Data.
- Las V's del Big Data.
- Big Data en los negocios.
- Ejemplos de Big Data.
- Tendencias del mercado.
- Azure HDInsight.

2 Exploración del Big Data

- Vista 360 del cliente.
- Análisis, diseño y operación.
- Procesamiento.
- Seguridad.
- Operatividad y consideraciones de implementación.
- Diseño de una solución.

3 Big Data 2.0

- Introducción a Spark.
- Propósitos de Spark.
- Componentes.
- Instalación.
- Configuración.
- Scala & Python.
- Herramientas de visualización de Big Data.

4 Big Data y ciencia de datos

- La plataforma del Big Data.
- El científico de Datos.
- El Proceso de la Ciencia de Datos.

5 Hadoop

- Ecosistema.
- Big Data y Cloud.
- Bases de datos NoSQL.
- Arquitectura Hadoop:
 - Componentes.
 - Cluster.
 - HDFS.
 - Despliegue.
- Componentes:
 - Map Reduce.
 - Pig
 - Hive
 - Flume.
 - Sqoop
 - Oozie.
- Administración:
 - Health Check.
 - Configuraciones.
 - Parámetros.
 - Buenas prácticas de administración.