

# Trabajo Final de Curso



Big Data y  
Análisis de  
Datos

## Tecnologías de la Información

## 1.TEMA

**Instalación de circuitos electroneumáticos de secuencias compuestas**

## 2.OBJETIVO

Al finalizar el presente trabajo final el estudiante tendrá la competencia de comprender de manera sólida los fundamentos del Big Data y el análisis de datos, además de contar con las habilidades esenciales para abordar y gestionar conjuntos masivos de datos. Este curso también estimulará una comprensión crítica sobre la relevancia del Big Data en una variedad de sectores industriales, cumpliendo las normas técnicas, las normas de seguridad y salud en el trabajo, actuando de manera responsable con el medio ambiente.

## 3.CONSIDERACIONES

El trabajo final consiste en resolver el caso práctico presentado, utilizando como referencia el problema planteado y las preguntas guía proporcionadas para orientar el desarrollo.

Los participantes deberán fundamentar sus propuestas en los conocimientos adquiridos a lo largo del curso, aplicando lo aprendido en las tareas y operaciones descritas en los contenidos curriculares.

## 4.CASO PRÁCTICO

La empresa DataCorp, dedicada al análisis de tendencias de mercado en el sector minorista, ha experimentado un aumento exponencial en la cantidad de datos generados a partir de múltiples fuentes, incluyendo redes sociales, compras en línea, sensores IoT en tiendas físicas y registros de clientes. Actualmente, la empresa enfrenta dificultades en el almacenamiento, procesamiento y análisis eficiente de estos datos, lo que impacta la toma de decisiones estratégicas.

Entre los principales problemas identificados se encuentran:

- Falta de una infraestructura adecuada para el almacenamiento de datos masivos.
- Largos tiempos de procesamiento en análisis de datos.
- Dificultad para extraer información relevante y útil para la toma de decisiones.

- Preocupaciones sobre el cumplimiento normativo y la seguridad de los datos.

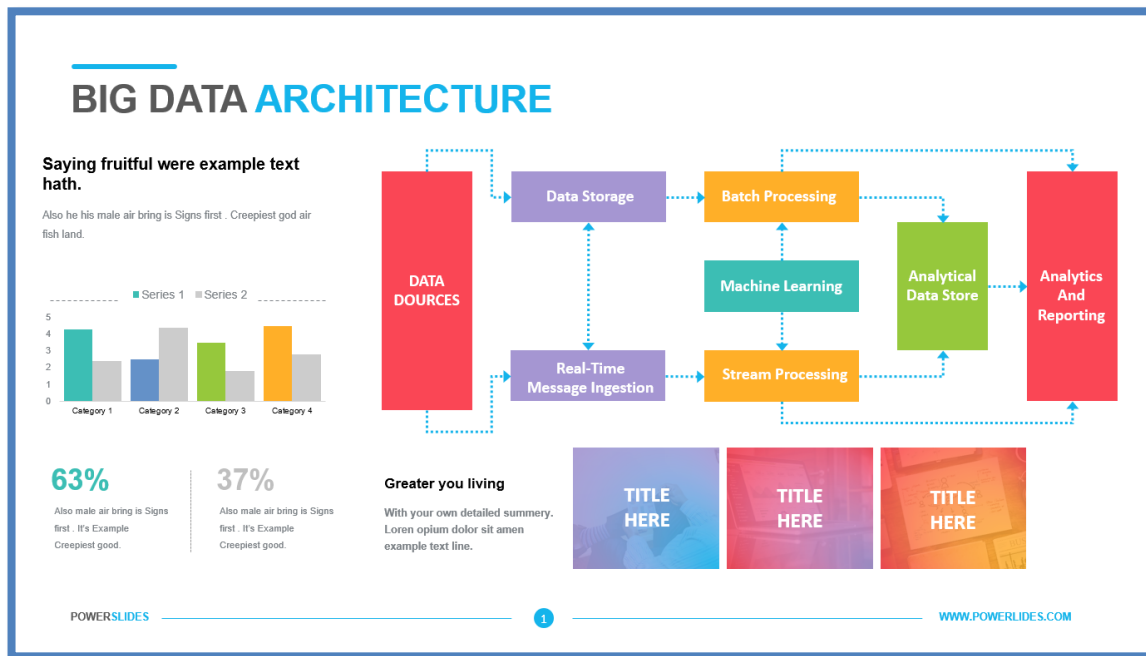
DataCorp busca implementar una solución de Big Data que le permita:

- Mejorar la capacidad de almacenamiento y acceso eficiente a los datos masivos generados.
- Implementar herramientas y tecnologías que optimicen el procesamiento de datos en tiempo real.
- Utilizar técnicas avanzadas de análisis para extraer información útil para el negocio.
- Garantizar la seguridad y cumplimiento normativo en la gestión de datos.
- Aplicar tecnologías emergentes para mejorar su competitividad en el mercado.

**La propuesta que permita solucionar el Caso Práctico debe incluir:**

- ✓ Explicación práctica de los conceptos de Big Data con ejemplos aplicados a la empresa.
- ✓ Identificación de las principales fuentes de generación de datos masivos dentro de DataCorp.
- ✓ Caso práctico donde se analice cómo el Big Data puede mejorar la toma de decisiones en la empresa.
- ✓ Describir la arquitectura Big Data que se utilizará en la solución.
- ✓ Comparación práctica entre sistemas de archivos distribuidos GFS y HDFS, aplicados a la empresa.
- ✓ Implementación de bases de datos NoSQL en un entorno práctico (MongoDB, Apache Cassandra, etc.).
- ✓ Investigación aplicada sobre las opciones de almacenamiento en la nube y su viabilidad para DataCorp.
- ✓ Análisis comparativo de frameworks como Hadoop, Spark y Flink en función de las necesidades de DataCorp.
- ✓ Ejemplo práctico de un proceso MapReduce aplicado al análisis de tendencias de clientes.
- ✓ Análisis comparativo de frameworks como Hadoop, Spark y Flink en función de las necesidades de DataCorp.

- ✓ Ejemplo práctico de un proceso MapReduce aplicado al análisis de tendencias de clientes.
- ✓ Presentar informe técnico detallando los procesos desarrollados.



### Arquitectura Big Data

## 5.PREGUNTAS GUÍAS

- **El desarrollo de las preguntas guías tienen el propósito de orientar la generación de su propuesta sobre el caso práctico.**
- 1) ¿Cuáles son las principales fuentes de datos masivos en la empresa y cómo pueden ser gestionadas de manera eficiente?
- 2) ¿Qué ventajas y desventajas tienen los diferentes sistemas de almacenamiento de datos masivos para la empresa?
- 3) ¿Cuál es el framework de procesamiento de datos más adecuado para los objetivos de DataCorp y por qué?
- 4) ¿Cómo se pueden aplicar herramientas de análisis de Big Data para mejorar la toma de decisiones empresariales en la empresa?
- 5) ¿Cuáles son los principales desafíos éticos y legales en la gestión de datos masivos dentro de DataCorp?

## 6.CONSIDERACIONES PARA EL ENTREGABLE

- ✓ Entregar una propuesta de solución para el caso práctico, fundamentado con los conocimientos adquiridos a lo largo del curso, aplicando lo aprendido con las tareas y operaciones descritas en los contenidos curriculares.
- ✓ Generar esquema y/o diagramas alineados a la propuesta de solución del caso práctico.

