



Gracias a



[openwebinars.net/cert/jOe7](https://openwebinars.net/cert/jOe7)



OpenWebinars certifica que

**ALBERTO JIMÉNEZ BONILLO**

Ha superado con éxito

**Curso de Implementación de una  
Data Pipeline con AWS EMR y  
PySpark**

Duración del curso

**3 horas**

Fecha de expedición

**13 septiembre 2024**

A handwritten signature in black ink, appearing to read 'Manuel Agudo', positioned over a large, light gray watermark of the OpenWebinars logo.

CEO de OpenWebinars

**Manuel Agudo**

## Contenido

# Curso de Implementación de una Data Pipeline con AWS EMR y PySpark

## 1. INTRODUCCIÓN

Presentación

## 2. EMR: CONCEPTOS Y CONFIGURACIÓN INICIAL

¿Qué es AWS EMR?

Introducción a PySpark

Configurar y lanzar el clúster de EMR

Test autoevaluación

## 3. ADMINISTRANDO EL CLUSTER DE EMR

Administrar el clúster de EMR

Lanzar una aplicación y un job con EMR Serverless

Configurar la seguridad del clúster

Monitorizar EMR Jobs

Test autoevaluación

## 4. CARGANDO LOS DATOS A ATHENA

¿Qué es AWS Athena?

Cargar datos de EMR a Athena

Diseños de casos reales con AWS EMR

Test autoevaluación

## 5. CONCLUSIONES

Conclusiones