



OpenWebinars certifica que

## **ANA ROA MEDINA**

Ha superado con éxito

Curso de Spark SQL: Programación de Scripts

Duración del curso

Fecha de expedición

5 horas

31 julio 2021

CEO de OpenWebinars

Manuel Agudo

### Contenido

# Curso de Spark SQL: Programación de Scripts

#### 1. INTRODUCCIÓN

Presentación

¿Qué es Apache Spark?

Modelo de programación

#### 2. FUNDAMENTOS DE PROGRAMACIÓN

Configuración del entorno en Windows

Configuración del entorno en Linux

Flujo de tratamiento

Ejemplos de lectura y escritura

Diferencia entre las tres APIs: DataFrame, Dataset y Case Class

Agregaciones

#### 3. SUPUESTO PRÁCTICO

Descripción del supuesto práctico

Configuración del proyecto

Agregaciones y Joins

UDFs y filtros

Librería externa con Java

Inclusión en el proyecto Scala de Spark

Creación y submisión de un jar con las tareas a ejecutar