



OpenWebinars certifica que  
**Joseba Gomez Carrasco**

Ha superado con éxito  
**Curso de Spark SQL:  
Programación de Scripts**

Duración del curso

**5 horas**

Fecha de expedición

**21 mayo 2020**

A handwritten signature in black ink, likely belonging to Manuel Agudo, the CEO of OpenWebinars.

CEO de OpenWebinars

**Manuel Agudo**

# 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