

Catálogo de Especialidades Formativas

PROGRAMA FORMATIVO

DESARROLLADOR SPARK BIG DATA CLOUDERA

Versión 3

Febrero 2026

GOBIERNO DE ESPAÑA
MINISTERIO DE TRABAJO Y ECONOMÍA SOCIAL
SERVICIO PÚBLICO DE EMPLEO ESTATAL
SEPE

SISTEMA NACIONAL DE EMPLEO

IDENTIFICACIÓN DE LA ESPECIALIDAD Y PARÁMETROS DEL CONTEXTO FORMATIVO

Denominación de la especialidad:	DESARROLLADOR SPARK BIG DATA CLOUDERA
Familia Profesional:	INFORMÁTICA Y COMUNICACIONES
Área Profesional:	DESARROLLO
Código:	IFCD25
Nivel de cualificación profesional:	3

Objetivo general

Dominio de Cloudera Data Platform (CDP) para diseñar e implementar aplicaciones que permitan ingestar, procesar y almacenar datos en un clúster Hadoop utilizando Apache Spark.

Relación de módulos de formación

Módulo 1	Introducción a SQL	10 horas
Módulo 2	Programación Python	50 horas
Módulo 3	Programación Scala	50 horas
Módulo 4	DENG-254: Preparing with Cloudera Data Engineering and Apache Spark	90 horas

Modalidades de impartición

Presencial

Duración de la formación

Duración total	200 horas

Requisitos de acceso del alumnado

Acreditaciones / titulaciones	Cumplir como mínimo alguno de los siguientes requisitos: -Título de Bachiller o equivalente -Título de Técnico Superior (FP de Grado superior) o equivalente -Haber superado la prueba de acceso a ciclos formativos de Grado Superior -Haber superado cualquier prueba oficial de acceso a la universidad -Certificado profesional de nivel 3 -Título de Grado Universitario o equivalente -Título de Master Universitario o equivalente
Experiencia profesional	No se requiere
Otros	Se recomiendan los siguientes requisitos mínimos: - Conocimientos generales de sistemas operativos, uso

Otros	básico de comandos Linux, Shell Scripting de Linux y programación en Shell script de Linux. - Conocimientos básicos de SQL y de algún lenguaje de programación funcional basada en objetos. - Inglés a nivel de lectura. Cuando el aspirante al curso no posea el nivel académico indicado, demostrará conocimientos suficientes a través de una prueba de acceso.

Justificación de los requisitos del alumnado

Para acreditar los requisitos de titulación, se documentará con la presentación de los títulos o certificados correspondientes.
-En caso de no disponer de nivel académico mínimo indicado, el aspirante demostrará conocimientos y competencias suficientes para participar en el curso con aprovechamiento mediante la superación de una prueba de acceso de competencias básicas correspondiente a la especialidad.
Además, el Centro comprobará mediante una prueba de acceso que el candidato posee conocimientos generales relacionados con la especialidad indicados anteriormente.

Prescripciones de formadores y tutores

Acreditación requerida	Cumplir como mínimo alguno de los siguientes requisitos: Cumplir como mínimo alguno de los siguientes requisitos: - Licenciado/a, Ingeniero/a, Arquitecto/a o el Título de Grado correspondiente u otras titulaciones equivalentes en los ámbitos de conocimiento de ingeniería, informática y de sistemas, ingeniería eléctrica, ingeniería electrónica, ingeniería de la telecomunicación, ingeniería industrial, matemáticas, estadística. - Diplomado/a, Ingeniero/a Técnico/a, Arquitecto/a Técnico/a o el Título de Grado correspondiente u otras titulaciones equivalentes en los ámbitos de conocimiento de ingeniería, informática y de sistemas, ingeniería eléctrica, ingeniería electrónica, ingeniería de la telecomunicación, ingeniería industrial, matemáticas, estadística. - Técnico superior de las familias profesionales: Informática y Comunicaciones, Electricidad y Electrónica
Experiencia profesional mínima requerida	Al menos 1 año en ocupaciones relacionadas con la especialidad.
Competencia docente	Se requiere acreditar un mínimo de 500 horas de experiencia docente o 250 horas si la experiencia está relacionada con especialidades o certificados de profesionalidad de la misma familia profesional en la que está encuadrada la presente especialidad. O estar en posesión de alguna de las siguientes titulaciones: -Certificado de Profesionalidad de Docencia de la Formación Profesional para el empleo (SSCE0110) o equivalente. -Certificado de Aptitud Pedagógica (CAP). -Título de licenciado en Pedagogía, Psicopedagogía o de Maestro en cualquiera de sus especialidades, graduado

Competencia docente	universitario en el ámbito de la Psicología o de la Pedagogía, o un título universitario oficial de posgrado en los citados ámbitos. -Título profesional de Especialización Didáctica, el Certificado de Cualificación Pedagógica, el Máster Universitario habilitante para el ejercicio de las Profesiones Reguladas de Profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Escuelas Oficiales de idiomas.
Otros	El formador deberá estar homologado como instructor en la correspondiente tecnología específica del fabricante. Excepcionalmente, cuando el profesor no acredite el requisito de titulación académica requerido, el centro formativo podrá valorar la posibilidad de acreditarlo mediante la emisión por parte del fabricante de una carta de validación como docente

Justificación de las prescripciones de formadores y tutores

Para acreditar tanto los requisitos de formación como de experiencia impartiendo formación en las horas mínimas estipuladas, se documentará con la presentación de títulos y certificados oficiales, vida laboral, certificados de empresa o contratos más liquidación (que demuestre que se han ejecutado).
Además, deberá disponer de la alguna de las siguientes certificaciones vigentes a la fecha de comienzo de la formación:
• Instructor de Cloudera.

Requisitos mínimos de espacios, instalaciones y equipamientos

Espacios formativos	Superficie m² para 15 participantes	Incremento Superficie/ participante (Máximo 30 participantes)
Aula de informática	45.0 m²	2.4 m² / participante

Espacio formativo	Equipamiento
Aula de informática	- Mesa y silla para el formador. - Mesas y sillas para el alumnado. - Material de aula. - Pizarra. - Impresora láser con conexión a red. - Pantalla y cañón de proyección. - PC instalado en red con posibilidad de impresión de documentos, cañón con proyección e Internet para el formador. La configuración mínima de los equipos que se recomienda es: CPU de 64 bits Intel Core i5 de novena generación o superior, o AMD similar; 1 disco duro SSD de 512GB o superior, máquina virtual en remoto con memoria RAM 104 GB, 2 discos SSD y procesador Intel AMD (24 cores); Pantalla 21´o superior, resolución de pantalla:1600*1080 para interfaces de gráficas de usuario, Gigabit Ethernet, soporte USB3 , teclado multimedia USB, ratón sensor óptico USB de 2 botones y rueda de desplazamiento.

	- PCs instalados en red e Internet con posibilidad de impresión para los alumnos. La configuración mínima de los equipos que se recomienda para cada alumno es: CPU de 64 bits Intel Core i5 de novena generación o superior, o AMD similar; 1 disco duro SSD de 512GB o superior, máquina virtual en remoto con memoria RAM 104GB, 2 discos SSD y procesador Intel AMD (24 cores); Pantalla 21´o superior, resolución de pantalla: 1600*1080 para interfaces de gráficas de usuario, Gigabit Ethernet, soporte USB3 , teclado multimedia USB, ratón sensor óptico USB de 2 botones y rueda de desplazamiento. - Software específico para el aprendizaje de cada acción formativa: o Licencia sistema operativo o Licencia del software antivirus o Licencias del software y herramientas necesarias para la impartición del curso (versión actualizada) aportado por el fabricante. o Acceso a los sistemas oficiales de Cloudera configurados específicamente con los ejercicios prácticos del curso aportados por el fabricante

La superficie de los espacios e instalaciones estarán en función de su tipología y del número de participantes. Tendrán como mínimo los metros cuadrados que se indican para 15 participantes y el equipamiento suficiente para los mismos.

En el caso de que aumente el número de participantes, hasta un máximo de 30, la superficie de las aulas se incrementará proporcionalmente (según se indica en la tabla en lo relativo a m²/participante) y el equipamiento estará en consonancia con dicho aumento.

No debe interpretarse que los diversos espacios formativos identificados deban diferenciarse necesariamente mediante cerramientos.

Las instalaciones y equipamientos deberán cumplir con la normativa industrial e higiénico-sanitaria correspondiente y responderán a medidas de accesibilidad y seguridad de los participantes.

En el caso de que la formación se dirija a personas con discapacidad se realizarán las adaptaciones y los ajustes razonables para asegurar su participación en condiciones de igualdad.

Aula virtual

Características
• La impartición de la formación mediante aula virtual se ha de estructurar y organizar de forma que se garantice en todo momento que exista conectividad sincronizada entre las personas formadoras y el alumnado participante así como bidireccionalidad en las comunicaciones. • Se deberá contar con un registro de conexiones generado por la aplicación del aula virtual en que se identifique, para cada acción formativa desarrollada a través de este medio, las personas participantes

en el aula, así como sus fechas y tiempos de conexión.

Vinculaciones con capacitaciones profesionales

Cloudera Data Engineer

Ocupaciones y puestos de trabajo relacionados

- 27111019 ANALISTAS DE SISTEMAS, NIVEL MEDIO (JUNIOR)
- 27111028 ANALISTAS DE SISTEMAS, NIVEL SUPERIOR (SENIOR)
- 27111037 INGENIEROS INFORMÁTICOS
- 27191013 AUDITORES-ASESORES INFORMÁTICOS
- 27191022 INGENIEROS TÉCNICOS EN INFORMÁTICA, EN GENERAL
- 38201017 PROGRAMADORES DE APLICACIONES INFORMÁTICAS

Requisitos oficiales de las entidades o centros de formación

Estar inscrito en el Registro de entidades de formación (Servicios Públicos de Empleo).

Los centros impartidores de formación Oficial de Cloudera deben cumplir el requisito: Acreditación de encontrarse autorizado por Cloudera Authorized Training Center de Cloudera.
Los requisitos mínimos tanto Hardware como Software serán los que marque el fabricante como recomendados en cada momento para las versiones actualizadas.
A los alumnos se les proporcionará la documentación oficial necesaria para el seguimiento del curso.

MÓDULO DE FORMACIÓN 1: Introducción a SQL

OBJETIVO

Identificar los conceptos clave y dominar las operaciones fundamentales de interacción con bases de datos relacionales utilizando el lenguaje SQL.

DURACIÓN:

10 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Identificación de conceptos clave de SQL
  - Definición de bases de datos y su importancia
  - Presentación del lenguaje SQL y sus usos comunes
- Manejo de operaciones fundamentales
  - Realización de consultas
  - Obtención, ordenación y filtrado de datos específicos de una tabla
  - Uso de operadores y funciones
  - Resumen y agrupación de datos
  - Creación, alteración y borrado de tablas
- Relaciones en BBDD y utilización de Joins.
  - Relaciones entre tablas.
  - Tipos y Operaciones con Joins para combinar datos de múltiples tablas.
  - Uso y aplicación de subconsultas
- Optimización de Consultas
  - Creación y uso de índices para mejorar el rendimiento de las consultas.
  - Creación y uso de vistas para simplificar y agilizar consultas.

Habilidades de gestión, personales y sociales

- Concienciación de la importancia y del uso de SQL
- Fomento del análisis crítico y la atención focalizada a la hora de analizar datos.

MÓDULO DE FORMACIÓN 2: Programación Python

OBJETIVO

Utilizar adecuadamente el lenguaje de programación Python para realizar las funciones esenciales relacionadas con objetos, colecciones de datos, control de flujo y librerías.

DURACIÓN:

50 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Identificación de los conceptos básicos de Python
- Conceptos básicos del lenguaje
- Sintaxis
- Definición y asignación de variables y tipos de datos
- Uso y manejo de variables
- Utilización de colecciones de datos
- Manipulación de datos complejos
- Uso de Herramientas de desarrollo para Python
- Manejo de estructuras de control de flujo
- Identificación de la estructura del programa
- Sentencias para el control de flujo de un programa
- Funcionalidades de las bibliotecas estándar de Python
- Importación y manejo de librerías
- Operaciones con expresiones matemáticas y expresiones regulares

Habilidades de gestión, personales y sociales

- Capacidad para obtener, procesar y asimilar nuevas habilidades y conocimientos en entornos cambiantes a la hora de programar.
- Habilidad para analizar correctamente la información disponible que posibilite la óptima gestión de procesos.
- Aptitud positiva y comprometida para la resolución de los problemas, identificando los componentes clave y las diversas maneras de abordaje.

MÓDULO DE FORMACIÓN 3: Programación Scala

OBJETIVO

Implementar soluciones basadas en programación funcional utilizando las funcionalidades del lenguaje Scala.

DURACIÓN:

50 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Identificación de los conceptos básicos de Scala.
- Introducción a los fundamentos Scala
- Sintaxis del lenguaje.
- Operaciones clave e implementación
- Uso de variables y tipos de datos en Scala
- Operaciones básicas de entrada/salida
- Agrupación de datos
- Control de flujo en Scala
- Manejo de estructuras complejas de datos (Tuples, List, Maps)
- Creación, uso e importación de librerías
- Gestión optimizada de datos

Habilidades de gestión, personales y sociales

- Concienciación de la importancia de los beneficios del lenguaje Scala
- Capacidad de adaptabilidad en la realización de acciones, bajo un modelo de flexibilidad para responder ante la transformación digital de las empresas.
- Aumento de la capacidad de análisis y toma de decisiones a la hora de implementar soluciones eficientes

MÓDULO DE FORMACIÓN 4:	DENG-254: Preparing with Cloudera Data Engineering and Apache Spark

OBJETIVO

Implementar y manejar las funcionalidades clave de Apache Spark en el desarrollo de aplicaciones distribuidas de alto rendimiento en la plataforma Cloudera Data Platform (CDP) utilizando Spark SQL, Hive y Hadoop (HDFS).

DURACIÓN:

90 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Introducción a HDFS
  - Visión general de HDFS.
  - Componentes e interacciones de HDFS.
  - Interacciones adicionales de HDFS.
  - Visión general de Ozone.
  - Ejercicio: Trabajar con HDFS.
- Introducción a YARN
  - Visión general de YARN.
  - Componentes e interacción de YARN.
  - Trabajar con YARN.
  - Ejercicio: Trabajar con YARN.
- Trabajar con RDDs
  - Conjuntos de datos distribuidos resilientes (RDDs).
  - Ejercicio: Trabajar con RDDs.
- Trabajar con DataFrames
  - Introducción a los DataFrames.
  - Ejercicio: Introducción a los DataFrames.
  - Ejercicio: Lectura y escritura de DataFrames.
  - Ejercicio: Trabajar con columnas.
  - Ejercicio: Trabajar con tipos complejos.
  - Ejercicio: Combinar y dividir DataFrames.
  - Ejercicio: Resumir y agrupar DataFrames.
  - Ejercicio: Trabajar con UDFs.
  - Ejercicio: Trabajar con ventanas.
- Introducción a Apache Hive
  - Acerca de Hive.
  - Transformación de datos con Hive QL.
- Trabajar con Apache Hive
  - Ejercicio: Trabajar con particiones.
  - Ejercicio: Trabajar con buckets.
  - Ejercicio: Trabajar con datos sesgados (Skew).
  - Ejercicio: Usar Serdes para digerir datos de texto.
  - Ejercicio: Usar tipos complejos para desnormalizar datos.
- Integración de Hive y Spark
  - Integración de Hive y Spark.
  - Ejercicio: Integración de Spark con Hive.
- Desafíos del procesamiento distribuido
  - Shuffle.
  - Skew.
  - Order.
- Procesamiento distribuido de Spark

- Procesamiento distribuido de Spark.
- Ejercicio: Explorar el orden de ejecución de consultas.

• Persistencia distribuida de Spark
- Persistencia de DataFrames y Datasets.
- Niveles de almacenamiento de persistencia.
- Visualización de RDDs persistentes.
- Ejercicio: Persistir DataFrames.

• Trabajar con datasets en Scala
- Trabajar con Datasets en Scala.
- Ejercicio: Usar Datasets en Scala.

Habilidades de gestión, personales y sociales

• Fomento de la autonomía y autoestima aplicando las metodologías estudiadas a la hora de abordar soluciones eficientes.
• Habilidad de pensamiento estratégico empresarial al alinear el uso de la tecnología para obtener un beneficio para la compañía.
• Habilidad para analizar correctamente la información disponible que posibilite la óptima gestión de procesos.

ORIENTACIONES METODOLÓGICAS

• La formación ha de ser eminentemente práctica.
• Los conceptos y contenidos a adquirir han de ir acompañados de ejemplos prácticos.
• El formador/a utilizará el método demostrativo que consiste en que 1º el formador/a muestra el uso de las funciones en la plataforma y 2º da tiempo a los alumnos para que ellos lo realicen después.
• Todas las unidades de aprendizaje tienen que ir acompañadas de ejercicios planteados por el profesorado, de los que después se mostrará la solución.
• La evaluación formativa o control de la comprensión durante la impartición es imprescindible para que los alumnos avancen eficazmente y el formador/a realice los ajustes necesarios, si fuera preciso.

EVALUACIÓN DEL APRENDIZAJE EN LA ACCIÓN FORMATIVA

La evaluación tendrá un carácter teórico-práctico y se realizará de forma sistemática y continua, durante el desarrollo de cada módulo y al final del curso.

Puede incluir una evaluación inicial de carácter diagnóstico para detectar el nivel de partida del alumnado.

La evaluación se llevará a cabo mediante los métodos e instrumentos más adecuados para comprobar los distintos resultados de aprendizaje, y que garanticen la fiabilidad y validez de la misma.

Cada instrumento de evaluación se acompañará de su correspondiente sistema de corrección y puntuación en el que se explicite, de forma clara e inequívoca, los criterios de medida para evaluar los resultados alcanzados por los participantes.

La puntuación final alcanzada se expresará en términos de Apto/ No Apto.

CERTIFICACIÓN DE FABRICANTE

La ejecución y financiación del programa formativo incluye la presentación de los

alumnos que han realizado el curso con aprovechamiento a los exámenes para obtener la certificación oficial del fabricante, que gestionará el centro y que en ningún caso supondrá coste alguno para el alumno.
En concreto, para esta acción formativa está incluido el siguiente examen de certificación oficial de Cloudera, o los que los sustituyan actualizados al momento de su impartición:
• Cloudera Data Engineer